ICSMD

International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD 2020)

October 15-17, 2020, Empark Grand Hotel, Xi' an, China
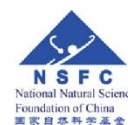

2020 CONFERENCE PROCEEDINGS

icsmd2020.aconf.org

TECHNICAL CO—SPONSERS

· IEEE IMS TC-7 Signals and Systems in Measurement
· IEEE IMS TC-3 Condition Monitoring and Fault Diagnosis Instrument

IEEE INSTRUMENTATION & MEASUREMENT SOCIETY®

FINANCIAL SUPPORTERS

National Natural Science Foundation of China

NSFC
National Natural Science Foundation of China
国家自然科学基金会

2020 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD2020) Proceedings

Table of Contents

Greetings from the General Chair

It's our great pleasure to invite you to join us for the 1st International Conference on Sensing, Measurement and Data Analytics in the Era of Artificial Intelligence (ICSMD 2020), which will invite experts and scholars in the vibrant field to meet and exchange ideas on the development of sensing methodologies, measurement technologies, and data analytics approaches with applications in various engineering domains in the era of artificial intelligence.

The ICSMD 2020 will be held in Xi'an, China on October 15-17, 2020. It is technically supported by the IEEE IMS TC-7 Signals and Systems in Measurement, the IEEE IMS TC-3 Condition Monitoring and Fault Diagnosis Instrument and the IEEE Instrumentation and Measurement Society, and financially sponsored by National Natural Science Foundation of China, and it will be organized by Xi'an Jiaotong University.

The ICSMD2020 will be the 1st edition of the conference and beginning of the series. Nowadays, sensing is the basis of intelligence, and measurement is the premise of intelligence. With the rapid development of intelligent manufacturing and intelligent equipment, there is an urgent need for the support of the basic theory and technology of sensing and measurement. Highly integrated micro/nano sensor, high-speed and high-precision measurement, digital twin and big data analysis, fault detection and isolation, intelligent health management, and so on, are more closely related and reinforced each other, which is exactly the aim of ICSMD 2020.

Xi'an is one of the oldest cities in China, with a history of more than 3,100 years. Called Chang'an before the Ming Dynasty, Xi'an is one of the birthplaces of the ancient civilization in the Yellow River Basin area of the country. This city used to be the capital city of 13 dynasties, and Xi'an is also the eastern terminus of the Silk Road and home to the Terracotta Warriors.

We are looking forward to meeting you in the very beautiful city Xi'an, China, during October 15-17, 2020. Definitely ICSMD 2020 will provide you a pleasant experience, new contacts and happy stay in Xi'an.

Ruqiang Yan

Chairman of ICSMD 2020

Organization Committee

Honory Chairs:
Prof. Asoke Nandi   Brunel University London, UK
Prof. Xuefeng Chen   Xi'an Jiaotong University, China

General Chairs:
Prof. Ruqiang Yan   Xi'an Jiaotong University, China
Dr. George Xiao   National Research Council, Canada

Technical Program Co-Chairs:
Prof. Datong Liu   Harbin Institute of Technology, China
Prof. Weihua Li   South China University of Technology, China

Special Session Co-Chairs:
Prof. Qingbo He   Shanghai Jiaotong University, China
Prof. Jinxing Liang   Southeast University, China

Publication Chair:
Dr. Yu Chen   Xi'an Jiaotong University, China

Publicity Chair:
Lingli Cui, Beijing University of Technology, China

Local Arrangement Chair:
Prof. Liuyang Zhang   Xi'an Jiaotong University, China

ICSMD 2020 Special Session Chairs

Session 1: Advanced sensing and intelligent computation for medical signals
Chengyu Liu （Southeast University ）
Zhi Tao （Soochow University）

Session 2: Fault prognosis and life prediction
Yi Qin （Chongqing University）
Chuan Li （Chongqing Technology and Business University）

Session 3: Advanced sensors and intelligent signal processing methods for navigation
Yuyong Xiong （Shanghai Jiao Tong University）
Dong Wang （Shanghai Jiao Tong University）

Session 4: Intelligent Sensing, Measurement & Data Analytics for Vehicles
Zhongkui Zhu （Soochow University）
Zaigang Chen （Southwest Jiaotong University）

Session 5: Condition Detection and Evaluation, System Control and Protection, Data Mining of Railway Track-vehicle-grid System
Zhigang Liu （Southwest Jiaotong University）
Alfredo Núñez （Delft University of Technology）

Session 6: Advanced sensors and intelligent signal processing methods for navigation
Haoqian Huang （Hohai University）
Di Liu （Southeast University）

Session7: Intelligent sensing and data analytics for smart manufacturing & Others
Yongbo Li （Northwestern Polytechnical University）
Siliang Lu （Anhui University）

Session 8: Flexible sensing and intelligent diagnosis/prognosis for rolling bearings & Session 15 NDT&E and Intelligent Monitoring & Others
Hongrui Cao （Xi'an Jiaotong University）
Yuhua Cheng （University of Electronic Science

and Technology of China）

Session 9: Quartz MEMS devices and technologies
Jing Ji (Xidian University)

Session10: Advanced Measurement and Data Processing for Aerospace
Jingli Yang （Harbin Institute of Technology）
Lianlei Lin （Harbin Institute of Technology）

Session 11: Intelligent Sensing and High Precision Measurement for Aerospace Advanced Sensing, Monitoring and Diagnosis in Smart Grid
Yu Chen （Xi'an Jiaotong University）
Zhe Li （Shanghai Jiaotong University）

Session 12: Intelligent Sensing and High Precision Measurement for Aerospace
Chairs: Guangcun Shan(Beihang University)**, Yong Zhang**(Xi'an Jiaotong University)

Session 13: Intelligent Anomaly Detection, Fault Diagnosis and Prognostics for Aero-engines
Jianzhoong Sun （Nanjing University of Aeronautics and Astronautics）
Liansheng Liu （Harbin Institute of Technology）

Session 14: THz Testing and Intelligent Sensing & Others
Liuyang Zhang （Xi'an Jiaotong University）
Shuncong Zhong （Fuzhou University）

Session 15: NDT&E and Intelligent Monitoring
Yuhua Cheng （University of Electronic Science and Technology of China）
Liuyang Zhang （Xi'an Jiaotong University）

Poster Presentations
Datong Liu （Harbin Institute of Technology）
Yu Chen （Xi'an Jiaotong University）

Social Events

Opening Ceremony
Date & Time: 08:40 – 9:00, October 16th (Friday)

Location: Main Hall, Empark Grand Hotel (3F)

Meals
All meals are included with your registration fee.

Banquet
Date & Time: 18:30 – 20:00, October 16th (Friday)

Location: Main Hall, Empark Grand Hotel (3F)

Closing Ceremony
Excellent Paper Awards will be awarded at Closing Ceremony. All presenters are advised to attend the closing ceremony.

Date & Time: 17:20 – 17:40, October 17th (Saturday)

Location: Main Hall, Empark Grand Hotel (3F)

ICSMD 2020 Reviewers

Liu Datong, Harbin Institute of Technology
Li Weihua, South China University of Technology
Libing Bai, University of Electronic Science and Technology of China
Zhipeng Cai, Southeast University
Zheng Cao, Anhui University
Hui Cao, Xi`an Jiaotong University
Huiliang Cao, North University of China
Zaigang Chen, Southwest Jiaotong University
Shiqian Chen, Southwest Jiaotong University
Yu Chen, Xi`an Jiaotong University
Cong Chen, University of Electronic Science and Technology of China
Zhixiong Chen, Shanghai University of Engineering Science
Xiaoxi Ding, Chongqing University
Xiaohui Gu, Shijiazhuang Tiedao University
Weiguo Huang, Soochow University
Haoqian Huang, Hohai University
Zhiyi He, Hunan University
Changbo He, Anhui University
Wangpeng He, Xidian University
Yawei Hu, Anhui University
Zhiwei Han, Southwest Jiaotong University
Hongkai Jiang, NorthwesternPolytechnicalUniversity
Jing Ji, Xidian University
Gaoqiang Kang, Southwest Jiaotong University
Xin Li, Hunan University
Xin Li, Nanjing Institute of Technology
Zhe Li, Shanghai Jiao Tong University
Yuwen Li, Southeast University
Yuanyuan Li, University of Electronic Science and Technology of China
Fucai Li, Shanghai Jiao Tong University
Hongkun Li, Dalian University of Technology
Naipeng Li, Xi`an Jiaotong University
Xuefeng Li, Tongji University
Li Li, Harbin University of Science and Technology
Zhixiong Li, Ocean University of China
Yongbin Liu, Anhui University
Fang Liu, Anhui University
Chengyu Liu, Southeast University
Wenqiang Liu, Southwest Jiaotong University
Jing Liu, Northwestern Polytechnical University
Pengpeng Liu, Unit 92942, People's Liberation Army
Siliang Lu, Anhui University
Jinxing Liang, Southeast University
Lei Mao, University of Science and Technology of China
Jiale Mao, Xi`an Jiaotong University
Jinhua Mi, University of Electronic Science and Technology of China
Yizhen Peng,Chongqing University
Fei Qin, University of Chinese Academy of Sciences
Na Qin, Southwest Jiaotong University
Zheng Qian, Beihang University
Yi Qin, Chongqing University

Guangcun Shan, Beihang University
Changqing Shen, Soochow University
Juanjuan Shi, Soochow University
Chong Shen, North University of China
Haidong Shao, Hunan University
Jinliang Shao, University of Electronic Science and Technology of China
Chao Sun, Harbin Institute of Technology
Lulu Tian, University of Electronic Science and Technology of China
Yunpu Wu, Southwest Jiaotong University
Haotian Wu, SouthChinaUniversityofTechnology
Dong Wang, Shanghai Jiao Tong University
Jun Wang, Soochow University
Xingxing Wang, Soochow University
Chenguang Wang, North University of China
Yi Wang, Chongqing University
Shuang Wang, Xi`an Jiaotong University
Weize Wang, East China University of Science and Technology
Hongrui Wang, Southwest Jiaotong University
Yanxue Wang, Beijing University of Civil Engineering and Architecture
Yu Wang, Xi`an Jiaotong University
Min Xia, Lancaster University
Wanli Xu, JiMei University
Yonggang Xu, Beijing University of Technology
Yuyong Xiong, Shanghai Jiao Tong University
Jiawei Xiang, Wenzhou University
Chang Yan, Southeast University
Haizi Yao, HuangHuai University
Wennian Yu, Chongqing University
Cheng Yang,China Shipbuilding Industry Research Institute of Integrated Technology and economy
Zhongkui Zhu, Soochow University
Jie Zhou, Hunan University
Meng Zhao, Xidian University
Liuyang Zhang, Xi`an Jiaotong University
Shuncong Zhong, Fuzhou University
Junping Zhong, Southwest Jiaotong University
Jie Zhang, University of Electronic Science and Technology of China
Ruiheng Zhang, University of Electronic Science and Technology of China

Keynote Speeches

Big Data Analytics for Intelligent Sensing and Measurement

Robert X. Gao, Ph.D.

Cady Staley Professor and Chair
Department of Mechanical and Aerospace Engineering
Case Western Reserve University
Cleveland, OH 44106-7222， USA
Email: robert.gao@case.edu

Abstract

Continued advancement in sensing and measurement technologies has led to an ever-increasing amount of data of a broad variety of forms and physical natures to be acquired from virtually all aspects of industrial and commercial fields. As rich information are embedded within these "big data", how to efficiently leverage them by means of effective data analytic methods to enhance manufacturing and contribute to economic development has become both a challenge and an opportunity.

This talk presents essential elements of and promising solutions enabled by big data analytics that complement measurement systems in the interpretation of high volume, broad variety, and low veracity data through enhanced pattern recognition and information extraction, with applications in machinery fault diagnosis, service life prognosis, and product quality control, to ultimately contribute to value creation. Case studies of machine learning methods such as deep learning in analyzing time series and image data and revealing mechanisms underlying manufacturing processes are discussed. Using assembly in manufacturing as a scenario, the talk highlights how multiphysics sensing and data analytics can be integrated for the recognition of current and prediction of future human actions during assembly operations to realize human-robot collaboration (HRC) in smart factories of the future.

Biographical Sketch

Dr. Gao is the Cady Staley Professor of Engineering and Department Chair of Mechanical and Aerospace Engineering at Case Western Reserve University in Cleveland, Ohio.   Since receiving his Ph.D. degree from the Technical University of Berlin, Germany in 1991, he has been working on multi-physics sensing, design and modeling of instrument systems, and machine learning techniques for improving the observability of dynamical systems such as manufacturing equipment and processes.   Dr. Gao is a Fellow of the Institute of Electrical and Electronic Engineers (IEEE), American Society of Mechanical Engineers (ASME), International Academy for Production Engineering (CIRP), and Society of Manufacturing Engineers (SME). He currently serves as a Senior Editor for the IEEE/ASME Transactions on Mechatronics, and is a recipient of the IEEE Best Application in Instrumentation and Measurement Award, IEEE Instrumentation and Measurement Society Technical Award, ASME Blackall Machine Tool and Gage Award, SME Eli Whitney Productivity Award, an NSF Early CAREER Award.

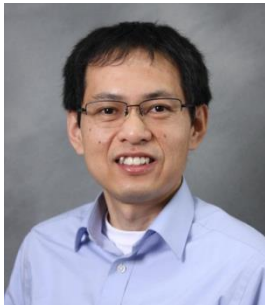# High-speed 3D Optical Sensing, Information Processing, and Applications

Song Zhang, Ph.D.

Professor and Assistant Head for Experiential Learning
School of Mechanical Engineering
Purdue University
West Lafayette, IN 47907, USA
Email: szhang15@purdue.edu

Abstract

Advances in optical imaging and machine/computer vision have provided integrated smart sensing systems for the manufacturing industry; and advanced 3D sensing could have profound impact on numerous fields, with broader applications including manufacturing, autonomous vehicles, and biomedical engineering. Our research addresses the challenges in high-speed, high-resolution 3D sensing and optical information processing. For example, we have developed a system that simultaneously captures, processes and displays 3D geometries at 30 Hz with over 300,000 measurement points per frame, which was unprecedented at that time (a decade ago). Our current research focuses on achieving speed breakthroughs by developing the binary defocusing techniques; and exploring novel means to store enormously large 3D data by innovating geometry/video compression methods. The binary defocusing methods coincide with the inherent operation mechanism of the digital-light-processing (DLP) technology, permitting tens of kHz 3D imaging speed at camera pixel spatial resolution. The novel methods of converting 3D data to regular 2D counterparts offer us the opportunity to leverage mature 2D data compression platform, achieving extremely high compression ratios without reinventing the whole data compression infrastructure. In this talk, I will present two platform technologies: 1) superfast 3D optical sensing; and 2) real-time 3D video communication. I will also cover some of the applications that we have been exploring including autonomous vehicles, biomedical engineering, forensic sciences, along with others.

Biographical Sketch

Song Zhang is a Professor and the Assistant Head for Experiential Learning, School of Mechanical Engineering at Purdue University. He received his Ph.D. (2005) and M.S. (2003) degrees in Mechanical Engineering from Stony Brook University, and B.S. (2000) degree from University of Science and Technology of China. His primary research focuses on high-speed 3D optical sensing/imaging and optical information processing. He has over 200 publications including 130 journal articles and 2 books. 16 of his journal articles were selected as cover page highlights. His publications have been cited over 10,900 times with an h-index of 51. Besides being utilized in academia, technologies developed by his team have been used by Radiohead (a rock band) to create a music video *House of Cards*; and by the law enforcement personnel to document crime scenes. He has received awards including AIAA Best Paper Award, IEEE ROBIO Best Conference Paper Award, Best of SIGGRAPH Disney Emerging Technologies Award, NSF CAREER Award, Stony Brook University's inaugural "Forty under 40 Alumni Award", Discovery in Mechanical Engineering award, CoE Early Career Faculty Research Excellence Award from Purdue and Iowa State University, Purdue University Faculty Scholar. He was a technical editor for IEEE/ASME Transactions on Mechatronics. He currently serves as an associate editor for Optics Express, as well as Optics and Lasers in Engineering. He is a fellow of SPIE and OSA.

Digital Twin: State-of-the-art and Its Application

Fei Tao, Ph.D., Professor

School of Automation Science and Electrical Engineering
Beihang University, P.R.China
Email: ftao@buaa.edu.cn

Abstract

The global academic research of digital twin (DT) is first investigated, and a comparative analysis of digital twin research in USA, Germany, and China is then given out. Ten industry applications of digital twin are then introduced, especially the application of digital twin shop-floor. In order to better understand and use digital twin, some hot topics related to digital twin will be discussed, such as the concept of digital twin, the applicable guideline of digital twin, standards of digital twin, and so on.

Biographical Sketch

Fei Tao is currently a Professor at the School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing, China. His current research interests are digital twin driven product design/manufacturing and service, and smart manufacturing service. In these fields, he has authored 4 monographs as the first author and published over 50 papers in Nature, CIRP Annals and IEEE/ASME Transactions, of which 20 are ESI high cited papers, and his publication has over 15000 citations in Google Scholar. Prof. Tao is a Global Highly Cited Researcher in 2019. He is currently the Editor-in-Chief of the International Journal of Service and Computing-Oriented Manufacturing (IJSCOM), the Associate Editor of Robotic and Computer Integrated Manufacturing (RCIM). He is also a CIRP Associate Member and IEEE Senior Member. He is currently the Vice-Dean of the Research Institute of Science and Technology of BUAA.

Invited Talk

How to Write and Publish a Scientific Paper in English

Fujun Liang, Ph.D

Abstract

The talk of Dr. Fujun Liang includes the following points:
1. What is scientific writing?
2. Deciding the paper style.
3. Collecting the writing materials.
4. Planning the structures.
5. Writing in formal language.
6. Focusing on scientific content.
7. Creating new value.
8. Preparing paper submission.

Biographical Sketch

Fujun Liang, member of the Democratic League of China, is currently a Senior Editor, quality director of CJME, part-time professor of Beijing University of Posts and Telecommunications, and National Science and Technology Expert. He teaches writing and editing skills for China Association for Science and Technology, China Editology Society of Science Periodicals, and China Machine Press. He has been invited to give lectures in universities, scientific research institutes and academic conferences for many times. He has published more than ten books, such as *Writing and Submission of SCI papers*, *Standard Writing and Editing of Academic Papers*, etc.

ICSMD 2020 Organizer

ICSMD 2020 Conference Sponsors

IEEE INSTRUMENTATION & MEASUREMENT SOCIETY

IEEE IMS TC-7 Signals and Systems in Measurement
IEEE IMS TC-3 Condition Monitoring and Fault Diagnosis Instrument

ICSMD 2020 Supporters

Iharbour Academy of Frontier Equipment

National Natural Science Foundation of China
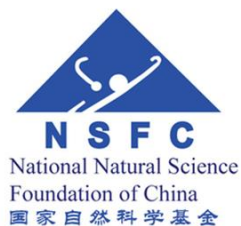
Awards

Best Paper Award

For the paper: Parameter Study of the Variable Reluctance Energy Harvester for Smart Railway Axle Box Bearing
Authors: Yun Gong, Sijia Wang, Zhengqiu Xie, Wenbin Huang
Chongqing University, China

For the paper: Modeling Voice Pathology Detection Using Imbalanced Learning
Authors: Ziqi Fan, Jinyang Qian, Di Wu, Yishen Xu, Baoyin Sun, Zhi Tao
Soochow University, China

Best Student Presentation Award

For the Paper: Decoupled Feature-Temporal CNN: Explaining Deep Learning-Based Machine Health Monitoring
Authors: Rui Zhao，Zhenghua Chen, Chaoyi Zhu, Zery Chan, Jinjiang Wang，Ruqiang Yan
Beihang University, China

Best Poster Award

For the paper: Modified Generative Adversarial Network for Super-Resolution of Terahertz Image
Authors: Zhen Zhang, Liuyang Zhang, Xuefeng Chen, Yafei Xu
Xi'an Jiaotong University, China

Technical Program

Oral Presentations
Friday, October 16, 2020

Keynote Speech 1   (9:00-9:45)

Big Data Analytics for Intelligent Sensing and Measurement

Robert X. Gao, Ph.D.

Cady Staley Professor and Chair
Department of Mechanical and Aerospace Engineering
Case Western Reserve University

Keynote Speech 2   (9:45-10:30)

High-speed 3D Optical Sensing, Information Processing, and Applications

Song Zhang, Ph.D.

Professor and Assistant Head for Experiential Learning
School of Mechanical Engineering
Purdue University

10:50-12:30
Session 4: Intelligent Sensing, Measurement & Data Analytics for Vehicles
Chairs: Zhongkui Zhu (Soochow University), Zaigang Chen (Southwest Jiaotong University)
Main Hall

Mengliu Pei (University of Science and Technology of China)

Chen Zhang (University of Science and Technology of China)

Mingruo Hu (Shanghai Jiao Tong University)

Lisa Jackson (Loughborough University)

Lei Mao (University of Science and Technology of China)

Shuyou Wu (Wuhan Second Ship Design and Research Institute)

Xu Peng (Wuhan Second Ship Design and Research Institute)

Lei Mao (University of Science and Technology of China)

Ao Liu (Wuhan Second Ship Design and Research Institute)

Liang Dong (Wuhan Second Ship Design and Research Institute)

Tong Wang (Wuhan Second Ship Design and Research Institute)

Chengliang Pan (HeFei University of Technology)

Shupeng Ren (HeFei University of Technology)

Fei Yang (HeFei University of Technology)

Chao Shi (HeFei University of Technology)

Heng Zuo (Nanjing Institute of Astronomical Optics & Technology, Chinese Academy of Sciences)

Zhihua Feng (University of Science and Technology of China)

(11:50-12:10) Vibration Feature of Locomotive Axle Box with a Localized Defect in its Bearing Outer Race ···17

Yuqing Liu (Southwest Jiaotong University)

Zaigang Chen (Southwest Jiaotong University)

(12:00-12:30) Buckling and Postbuckling Analysis Method of Stretchable and Flexible Sensor Networks Based on ABAQUS···21

Shuguang Hu (Nanjing University of Aeronautics and Astronautics)

Yu Wang (Nanjing University of Aeronautics and Astronautics)

Lei Qiu (Nanjing University of Aeronautics and Astronautics)

10:50-12:30
Session 6: Advanced Sensors and Intelligent Signal Processing Methods for Navigation
Chairs: Haoqian Huang (Hohai University), Di Liu (Southeast University)
International Hall

(10:50-11:10) Visual Inertial Calibration of Mobile Robotic System Based on Reinforcement Learning ···27

Wenxing Zhu (Southeast University)

Lihui Wang (Southeast University)

(11:10-11:30) Geomagnetic Matching Navigation for Aircraft Based on Improved Iterative Closest Contour Point Algorithm ···33

Ninghui Xu (Southeast University)

Lihui Wang (Southeast University)

(11:30-11:50) An Improved Interacting Multiple Model Algorithm for INS/DVL Integrated Navigation System ···39

Lanhua Hou（Southeast University）

Xiaosu Xu (Southeast University)

Yiqing Yao (Southeast University)

Di Wang (Southeat University)

(11:50-12:10) The Implementation and Comparison Between Kalman Filter-based and Vector Tracking Loops ··45

Zhe Yan (Southeast University)

Xiyuan Chen (Southeast University)

Xinhua Tang (Southeat University)

Xuefen Zhu (Southeat University)

10:50–12:30
Session 12: Intelligent Sensing and High Precision Measurement for Aerospace
Chairs: Guangcun Shan (Beihang University), Yong Zhang (Xi'an Jiaotong University)
No.7 Meeting Room

(10:50-11:10) Graph Theory Based Localization of Wireless Sensor Networks for Radio Irregularity Cases ···51

Xiaofeng Ma (Beihang University)

Ning Yu (Beihang University)

Tianle Zhou (Beihang University)

Renjian Feng (Beihang University)

Yinfeng Wu　(Beihang University)

10:50–12:30
Session 9: Quartz MEMS Devices and Technologies
Chairs: Jing Ji (Xidian University)
No.8 Meeting Room

Shiqi Hou (Tongji University)

Xuefeng Li (Tongji University)

14:00-15:40
Session 5: Condition Detection and Evaluation, System Control and Protection, Data Mining of Railway Track-vehicle-grid System
Chairs: Zhigang Liu(Southwest Jiaotong University), Alfredo Núñez(Delft University of Technology)
Main Hall

(14:00-14:20) Muscle Temperature Sensing and Control with A Wearable Device for Hand Rehabilitation of People After Stroke ···**94**

Han Li (Southeast University)

Jun Zhang (Southeast University)

Chaojun Jiang (Southeast University)

Qi Liu　(Southeast University)

Maozeng Zhang (Southeast University)

Jingsong Zhou (Southeast University)

(14:20-14:40) Action-driven Reinforcement Learning for Improving Localization of Brace Sleeve in Railway Catenary ···**100**

Junping Zhong (Southwest Jiaotong University)

Zhigang Liu (Southwest Jiaotong University)

Hongrui Wang (Delft University of Technology)

Wenqiang Liu (Southwest Jiaotong University)

Cheng Yang (Southwest Jiaotong University)

Alfredo Núñez (Delft University of Technology)

(14:40-15:00) An Automatic Defect Detection Method for Catenary Bracing Wire Components Using Deep Convolutional Neural Networks and Image Processing·······················**106**

Wenqiang Liu (Southwest Jiaotong University)

Dan Wang (Southwest Jiaotong University)

Cheng Yang (Southwest Jiaotong University)

Yuyang Li (Southwest Jiaotong University)

Hui Wang (Southwest Jiaotong University)

Zhigang Liu (Southwest Jiaotong University)

(15:00-15:20) Pantograph Vibration Interference Signal Recognition Based on SVM Classification ···**112**

Yuming Ding (China Academy of Railway Sciences)

Jun Zhao (China Academy of Railway Sciences Corporation Limited)

Jinzhao Liu (China Academy of Railway Sciences Corporation Limited)

Wenxuan Zhang (China Academy of Railway Sciences Corporation Limited)

Xiaodi Xu (China Academy of Railway Sciences Corporation Limited)

Zhipeng Yang (China Academy of Railway Sciences Corporation Limited)

(15:20-15:40) Coupling Deep Models and Extreme Value Theory for Open Set Fault Diagnosis ···**118**

Xiaolei Yu (Xi'an Jiaotong University)

Zhibin Zhao (Xi'an Jiaotong University)

Xingwu Zhang (Xi'an Jiaotong University)

Chuang Sun (Xi'an Jiaotong University)

Qiyang Zhang (Xi'an Jiaotong University)

Xuefeng Chen (Xi'an Jiaotong University)

14:00-15:40
Session 6: Advanced Sensors and Intelligent Signal Processing Methods for Navigation
Chairs: Haoqian Huang (Hohai University), Di Liu (Southeast University)
International Hall

Chengshuai Zhao (North University of China)

Xiaojie Liu (North University of China)

Chong Shen (North University of China)

Jun Tang (North University of China)

Jun Liu (North University of China)

Huiliang Cao (North University of China)

Rui Zhao (North University of China)

Qi Cai (North University of China)

Yunbo Shi (North University of China)

Li Liu (Jinzhou 777 Microelectronics Co., Ltd)

Haoqian Huang (Hohai University)

Jiacheng Tang (Hohai University)

Chao Wang (Hohai University)

Di Liu (Southeast University)

Xiyuan Chen (Southeast University)

Xiao Liu (Southeast University)

Zhe Yan (Southeast University)

Xianbin Sun (Qingdao University of Technology)

Xinming Jia (Qingdao University of Technology)

Yi Zheng (Shandong Academy of Sciences)

Zhen Wang (Shandong Academy of Sciences)

14:00-15:40
Session 12: Intelligent Sensing and High Precision Measurement for Aerospace
Chairs: Guangcun Shan (Beihang University), Yong Zhang (Xi'an Jiaotong University)
No. 7 Meeting Room

Yu Chen (Xi'an Jiaotong University)

Yan Yang (Xi'an Jiaotong University)

Guorui Huang (Xi'an Jiaotong University)

Hanzhi Li (Xi'an Jiaotong University)

Changxi Li (Xi'an Jiaotong University)

Shuang Wang (Xi'an Jiaotong University)

Yonghong Cheng (Xi'an Jiaotong University)

(14:20-14:40) Analysis of Secondary Electron Yield and Energy Spectrum of Metal Materials based on Furman Model

Zecai Chen (Xi'an Jiaotong University)

Yu Chen (Xi'an Jiaotong University)

Guorui Huang (Xi'an Jiaotong University)

Changxi Li (Xi'an Jiaotong University)

Qingyun Shi (Xi'an Jiaotong University)

Shuang Wang (Xi'an Jiaotong University)

(14:40-15:00) Theoretical Calculation of Photoemission Yield Spectrum of Space Metal Materials

Changxi Li (Xi'an Jiaotong University)

Yu Chen (Xi'an Jiaotong University)

Yan Yang (Xi'an Jiaotong University)

Zecai Chen (Xi'an Jiaotong University)

Hanzhi Li (Xi'an Jiaotong University)

Shuang Wang (Xi'an Jiaotong University)

(15:00-15:20) Development of a Measurement System for the Secondary Electron Emission Yield Spectrum of Space Materials

Yu Chen (Xi'an Jiaotong University)

Guorui Huang (Xi'an Jiaotong University)

Yan Yang (Xi'an Jiaotong University)

Qingyun Shi (Xi'an Jiaotong University)

Zecai Chen (Xi'an Jiaotong University)

Shuang Wang (Xi'an Jiaotong University)

Yonghong Cheng (Xi'an Jiaotong University)

(15:20-15:40) In-pipe Detection System Based on Magnetic Flux Leakage and Eddy Current Detection

Xiaolin Liu   (Beihang University)

Chun Hu (Beihang University)

Peng Peng (Chongqing Dexin Robot Testing Center Limited Company)

Rui Li (Petrochina Pipeline Company)

Xiaoming Zhao (Petrochina Pipeline Company)

Dezhi Zheng (Beihang University)

14:00-15:40
Session 11: Advanced Sensing, Monitoring and Diagnosis in Smart Grid
Chairs: Yu Chen (Xi'an Jiaotong University), Zhe Li(Shanghai Jiaotong University)
No. 8 Meeting Room

(14:00-14:20) Fault Accurate Location Method in UHV GIL Based on Transient Voltage Travelling Wave ···170

Yong Wang (Guangzhou Power Supply Co. Ltd Electric Power Test & Research Institute)

Dengwei Ding (Sichuan Energy Internet Research Institute, Tsinghua University)

Lu Zhu (Guangzhou Power Supply Co. Ltd Electric Power Test & Research Institute)

Wenxiong Mo (Guangzhou Power Supply Co. Ltd Electric Power Test & Research Institute)

Weinan Fan (Guangzhou Power Supply Co. Ltd Electric Power Test & Research Institute)

Ziwei Zhang (Sichuan Energy Internet Research Institute, Tsinghua University)

(14:20-14:40) An Optical Partial Discharge Localization Method Based on Simulation and Machine learning in GIL ···174

Yiming Zang (Shanghai Jiao Tong University)

Yong Qian (Shanghai Jiao Tong University)

Hui Wang (Shanghai Jiao Tong University)

Antian Xu (Fudan University)

Gehao Sheng (Shanghai Jiao Tong University)

Xiuchen Jiang (Shanghai Jiao Tong University)

(15:00-15:20) The 500kV Oil-filled Submarine Cable Temperature Monitoring System Based on BOTDA Distributed Optical Fiber Sensing Technology ····························180

Yu Chen (Xi'an Jiaotong University)

Shuang Wang (Xi'an Jiaotong University)

Yi Hao (Xi'an Jiaotong University)

Kai Yao (Xi'an Jiaotong University)

Hanzhi Li (Xi'an Jiaotong University)

Feng Jia (Xi'an Jiaotong University)

Qingyun Shi (Xi'an Jiaotong University)

Dongli Yue (Xi'an Jiaotong University)

Yonghong Cheng (Xi'an Jiaotong University)

(15:20-15:40) Stator Inter-turns Short Circuit Fault Detection in DFIG using Empirical Mode Decomposition Method on Leakage Flux ···································184

Attiq Ur Rehman (Xi'an Jiaotong University)

Yu Chen (Xi'an Jiaotong University)

Guorui Huang (Xi'an Jiaotong University)

Yan Yang (Xi'an Jiaotong University)

Shuang Wang (Xi'an Jiaotong University)

Yihan Zhao (Xi'an Jiaotong University)

Yong Zhao (Xi'an Thermal Power Research Institute Co. Ltd)

Yonghong Cheng (Xi'an Jiaotong University)

Toshikatsu Tanaka (Waseda University)

16:00-18:00
Session 2: Fault Prognosis and Life Prediction
Chairs: Yi Qin(Chongqing University), Chuan Li(Chongqing Technology and Business University)
Main Hall

(16:20-16:40) Principal Component Analysis Based Kullback-Leibler Divergence for Die Cracks Detection ···**224**

Sha Wei (Shanghai Jiao Tong University)

Dong Wang (Shanghai Jiao Tong University)

Zhike Peng (Shanghai Jiao Tong University)

(16:40-17:00) Research on Inductively Coupled Full Duplex Communication Method with Power Transmission···**229**

Tao Zhang (Chongqing university)

Deqi Zhang (Chongqing university)

Wenbin Huang (Chongqing university)

(17:00-17:20) Ultra-micro Vibration Measurement Method Using CW Doppler Radar··········**235**

Songxu Li (Shanghai Jiao Tong University)

Yuyong Xiong (Shanghai Jiao Tong University)

Zesheng Ren (Shanghai Jiao Tong University)

Changzhan Gu (Shanghai Jiao Tong University)

Zhike Peng (Shanghai Jiao Tong University)

(17:20-17:40) A Selective Anchor Node Method Based on Tetrahedral Volume for Three-dimensional DFL Node Self-localization ···**238**

Binghua Huang (Nanjing University of Science and Technology)

XuLiang Qin (NanJing University of Science and Technology)

JiaXing Yang (NanJing University of Science and Technology)

Manyi Wang (NanJing University of Science and Technology)

16:00-18:00
Session 13: Intelligent Anomaly Detection, Fault Diagnosis and Prognostics for Aero-engines
Chairs: Jianzhoong Sun (Nanjing University of Aeronautics and Astronautics), Liansheng Liu (Harbin Institute of Technology)
No.7 Meeting Room

(16:00-16:20) Enhancing the Reliability of the Quadrotor by Formulating the Control System Model ···**242**

Zhuo Zhi (Harbin Institute of Technology)

Liansheng Liu (Harbin Institute of Technology)

Datong Liu (Harbin Institute of Technology)

(16:20-16:40) An SOC and SOP Joint Estimation Method of Lithium-ion Batteries in Unmanned Aerial Vehicles ···**247**

Wanqing Cheng (Harbin Institute of Technology)

Zhiheng Yi (Harbin Institute of Technology)

Jun Liang (Harbin Institute of Technology)

Yuchen Song (Harbin Institute of Technology)

Datong Liu (Harbin Institute of Technology)

(16:40-17:00) UAV Anomaly Detection Using Active Learning and Improved S3VM Model······**253**

Dawei Pan (Harbin Engineering University)

Longqiang Nie (Harbin Engineering University)

Weixin Kang (Harbin Engineering University)

Zhe Song (Harbin Engineering University)

(17:00-17:20) Research on Multivariate Variational Mode Decomposition Method and Its Application to Bearing Fault Diagnosis ·······································259

Qiuyu Song (Soochow University)

Xingxing Jiang (Soochow University)

Jun Wang (Soochow University)

Changqing Shen (Soochow University)

Juanjuan Shi (Soochow University)

Weiguo Huang (Soochow University)

Zhongkui Zhu (Soochow University)

(17:20-17:40) Development and Application of Aero-engine Experimental Data Mining Algorithm Library··264

Zichen Yan (Nanjing University of Aeronautics and Astronautics)

Haonan Zhang   (University of Chinese Academy of Sciences)

Jianzhong Sun (Nanjing University of Aeronautics and Astronautics)

Yi Yang   (Nanjing University of Aeronautics and Astronautics)

Guangwei Xia   (AECC Hunan Aviation Powerplant Research Institute)

(17:40-18:00) Differentiable Architecture Search for Aeroengine Bevel Gear Fault Diagnosis ···270

Zheng Zhou (Xi'an Jiaotong University)

Tianfu Li (Xi'an Jiaotong University)

Zhibin Zhao (Xi'an Jiaotong University)

Chuang Sun (Xi'an Jiaotong University)

Ruqiang Yan (Xi'an Jiaotong University)

Xuefeng Chen (Xi'an Jiaotong University)

16:00-18:00
Session 8: Flexible sensing and intelligent diagnosis/prognosis for rolling bearings &
Session 15 NDT&E and Intelligent Monitoring & Others
Chairs: Hongrui Cao(Xi'an Jiaotong University), Yuhua Cheng(University of Electronic Science and Technology of China)
No.8 Meeting Room

(16:00-16:20) Rolling Bearing Fault Diagnosis based on Horizontal Visibility Graph and Graph Neural Networks··275

Chenyang Li (Southeast University)

Lingfei Mo (Southeast University)

Ruqiang Yan (Xi'an Jiaotong University)

(16:20-16:40) Decoupled Feature-Temporal CNN: Explaining Deep Learning-Based Machine Health Monitoring ···280

Rui Zhao (Harveston Asset Management)

Zhenghua Chen (Institute for Infocomm Research)

Chaoyi Zhu (Beihang University)

Zery Chan (Institute for Infocomm Research)

Jinjiang Wang (China University of Petroleum-Beijing)

Ruqiang Yan(Xi'an Jiaotong Univeristy)

Saturday, October 17, 2020

Keynote Speech 3   (9:00-9:50)
Digital Twin: State-of-the-art and Its Application
Fei Tao, Ph.D., Professor
School of Automation Science and Electrical Engineering
Beihang University, P.R.China

Lecture- How to Write and Publish a Scientific Paper in English (16:00-17:20)
Fujun Liang
Senior Editor, quality director of CJME, part-time professor of Beijing University of Posts and Telecommunications, and National Science and Technology Expert

10:10–12:10
Session 1: Advanced Sensing and Intelligent Computation for Medical Signals
Chairs: Chengyu Liu（Southeast University），Zhi Tao(Soochow University)
Main Hall

(10:10-10:30) A Bayesian Fusion Model for Heart Rate Annotations·····································307
Jianan Di (Southeast University)
Jianqing Li (Southeast University)
Chengyu Liu (Southeast University)

(10:30-10:50) Continuous Estimation of Left Ventricular Hemodynamic Parameters Based on Heart Sound and PPG Signals Using Deep Neural Network·····································313
Tengfei Feng (Dalian University of Technology)
Hong Tang (Dalian University of Technology)
Miao Wang (Dalian University of Technology)
Chi Zhang (Dalian University of Technology)
Hongkai Wang (Dalian University of Technology)
Fengyu Cong (Dalian University of Technology)

(10:50-11:10) Voice Pathology Detection and Multi-classification Using Machine Learning Classifiers·····································319
Yuanbo Wu (Soochow University)
Changwei Zhou (Soochow University)
Ziqi Fan (Soochow University)
Yihua Zhang (Soochow University)
Xaojun Zhang (Soochow University)
Zhi Tao (Soochow University)

(11:10-11:30) Classification of Normal and Pathological Voices Using Convolutional Neural Network·····································325
Changwei Zhou (Soochow University)
Lili Zhang (Soochow University)
Xiaojun Zhang   (Soochow University)
Yuanbo Wu (Soochow University)
Di Wu   (Soochow University)
Zhi Tao (Soochow University)

(11:30-11:50) Modeling Voice Pathology Detection Using Imbalanced Learning

Ziqi Fan (Soochow University)

Jinyang Qian (Soochow University)

Baoyin Sun (Soochow University)

Di Wu (Soochow University)

Yishen Xu (Soochow University)

Zhi Tao (Soochow University)

(11:50-12:10) Multi-label Feature Selection for Long-term Electrocardiogram Signals

Yuwen Li (Southeast University)

Zhimin Zhang (Science and Technology on Information Systems Engineering Laboratory)

Fan Zhou (Southeast University)

Yantao Xing (Southeast University)

Jianqing Li (Nanjing Medical University)

Chengyu Liu (Southeast University)

10:10–12:10
Session10: Advanced Measurement and Data Processing for Aerospace
Chairs: Jingli Yang (Harbin Institute of Technology), Lianlei Lin (Harbin Institute of Technology)
No.7 Meeting Room

(10:10-10:30) A Robotic Sensor Node for Mechanical Property Detection of Material on Asteroid Surface

Jun Zhang (*Southeast University*)

Maozeng Zhang (Southeast University)

Yizhuang Ding (Southeast University)

Liuchen Chen (Southeast University)

Qixuan Li (Southeast University)

Minghan Qin (Southeast University)

(10:30-10:50) Research and Verification of Multi-Satellite Thermal Vacuum Test Method

Xiaofeng Zhang (Innovation Academy for Microsatellites of CAS)

Li Wu (Innovation Academy for Microsatellites of CAS)

Hong Liu(Innovation Academy for Microsatellites of CAS)

Jianchao Feng (Innovation Academy for Microsatellites of CAS)

Meijuan Xu (Innovation Academy for Microsatellites of CAS)

Rui Cheng (Innovation Academy for Microsatellites of CAS)

(10:50-11:10) Satellite Control and Data Processing Unit Software Design based on Multi-core Processor

Junwang He (Shanghai Engineering Center for Microsatellites)

Luyang Xu(University of Science and Technology of China)

Dongxiao Xu (Shanghai Engineering Center for Microsatellites)

Shunjing Yu (Shanghai Engineering Center for Microsatellites)

Kaige Wang (School of Information Science and Engineering)

Liang Chang (Shanghai Engineering Center for Microsatellites)

(11:10-11:30) Analysis of Ionospheric Scintillation Detection based on Machine Learning

Mengying Lin (Southeast University)

Xuefen Zhu (Southeast University)

Yimei Luo (Southeast University)

Fan Yang (Southeast University)

(11:30-11:50) Detection Method of Solar Radio Bursts Based on Support Vector Machine Model ··**362**

Yimei Luo (Southeast University)

Xuefen Zhu (Southeast University)

Mengying Lin (Southeast University)

Fan Yang (Southeast University)

(11:50-12:10) Optimal Placement of Blade Tip Timing Sensors Considering Multi-mode Vibration Using Evolutionary Algorithms ··**367**

Jinghui Xu (Xi'an Jiaotong University)

Baijie Qiao (Xi'an Jiaotong University)

Zhibo Yang (Xi'an Jiaotong University)

Yuanchang Chen (University of Massachusetts Lowell)

Xuefeng Chen (Xi'an Jiaotong University)

10:10–12:10
Session 15: NDT&E and Intelligent Monitoring
Chairs: Yuhua Cheng (University of Electronic Science and Technology of China), Liuyang Zhang (Xi'an Jiaotong University)
No.10 Meeting Room

(10:10-10:30) Eddy Current Thermography for the Detection of Conductive Defects in Composite Insulators··**373**

Yanxin Tu (Tsinghua Shenzhen International Graduate school)

Chenjun Guo (Electric Power Research Institute of Yunnan Power Grid Co.)

Hongwei Mei (Tsinghua Shenzhen International Graduate school)

Lishuai Liu (Tsinghua Shenzhen International Graduate school)

Chenglong Cong (Tsinghua Shenzhen International Graduate school)

Liming Wang (Tsinghua Shenzhen International Graduate school)

(10:30-10:50) Model-Based Parameter Estimation Method for Terahertz Signals ··············**378**

Yafei Xu (Xi'an Jiaotong University)

Liuyang Zhang (Xi'an Jiaotong University)

Xuefeng chen (Xi'an Jiaotong University)

Zhen Zhang (Xi'an Jiaotong University)

Zhonglei Shen (Xi'an Jiaotong University)

Donghai Han (Xi'an Jiaotong University)

(10:50-11:10) Battery State of Health Estimation with Incremental Capacity Analysis Technique··**384**

Yuanyuan Li (University of Electronic Science and Technology of China)

Hanmin Sheng (University of Electronic Science and Technology of China)

Yuhua Cheng (University of Electronic Science and Technology of China)

(11:10-11:30) Convolutional LSTM Networks for Vibration-based Defect Identification of the Composite Structure ···390

Ruiheng Zhang (University of Electronic Science and Technology of China)

Rui Li (Shanghai Space Propulsion Technology Research Institute)

Libing Bai (University of Electronic Science Technology of China)

Lulu Tian (University of Electronic Science Technology of China)

Jie Zhang (University of Electronic Science Technology of China)

Zhen Liu (University of Electronic Science and Technology of China)

(11:30-11:50) Study on the Lowest Spatial Resolution of Magnetic Flux Leakage Testing for Weld Cracks··396

Chunrui Feng (University of Electronic Science Technology of China),

Zhen Zhang (Shanghai Space Propulsion Technology Research Institute)

Libing Bai (University of Electronic Science Technology of China)

Lulu Tian (University of Electronic Science Technology of China)

Jie Zhang (University of Electronic Science Technology of China)

Yuhua Cheng (University of Electronic Science and Technology of China)

(11:50-12:10) Application of Tensor Decomposition Methods in Eddy Current Pulsed Thermography Sequences Processing ··401

Yiping Liang (University of Electronic Science and Technology of China)

Libing Bai (University of Electronic Science and Technology of China)

Jinliang Shao (University of Electronic Science and Technology of China)

Yuhua Cheng (University of Electronic Science and Technology of China)

14:00–15:40
Session 1: Advanced Sensing and Intelligent Computation for Medical Signals & Others
Chairs: Chengyu Liu（Southeast University), Zhi Tao (Soochow University)
Main Hall

(14:00-14:20) A Portable NeuECG Monitoring System for Cardiac Sympathetic Nerve Activity Assessment···407

Yantao Xing (Southeast University)

Jianqing Li (Southeast University)

Zhengyuan Hu (Southeast University)

Yuwen Li   (Southeast University)

Yike Zhang (The First Affiliated Hospital of Nanjing Medical University)

Chang Cui (The First Affiliated Hospital of Nanjing Medical University)

Cheng Cai (The First Affiliated Hospital of Nanjing Medical University)

Chengyu Liu (Southeast University)

(14:20-14:40) An Octave Convolution Neural Network-based QRS Detector····················413

Wei Liu (Southeast University)

Xingyao Wang(Southeast University)

Hongxiang Gao (Southeast University)

Chenxi Yang (Southeast University)

Jianqing Li (Southeast University)

Chengyu Liu (Southeast University)

(14:40-15:00) Capacitance Measurement of Molten Metal Level in Continuous Casting System··**419**

Zhe Zhuang (Xi'an Jiaotong University)

Yu Zhang (Shanghai Hinner Information Technology Co., Ltd)

Tonghao Zhou (Xi'an Jiaotong University)

Shengchang Ji (Xi'an Jiaotong University)

(15:00-15:20) Application of Wire Mesh Sensor to Identify Gas-liquid Flow Patterns in a Horizontal Pipe ··**424**

Shuai Liu (Shanghai Jiao Tong University)

Li Liu (Shanghai Jiao Tong University)

Jiarong Zhang (Shanghai Jiao Tong University)

Hanyang Gu (Shanghai Jiao Tong University)

Qi Zhang (Shanghai Jiao Tong University)

(15:20-15:40) Flow Visualization of Centrifugal Pump by the Combination of LIF and PIV···**429**

Qi Zhang (Shanghai Jiao Tong University)

Hanyang Gu (Shanghai Jiao Tong University)

Shuai Liu(Shanghai Jiao Tong University)

Junlong Li (Shanghai Jiao Tong University)

Sichao Tan (Harbin Engineering University)

Jianke Su (support Center of Nuclear Technology, CAEA)

14:00-15:00
Session 14：THz Testing and Intelligent Sensing & Others
Chairs: Liuyang Zhang (Xi'an Jiaotong University), Shuncong Zhong (Fuzhou University)
No.7 Meeting Room

(14:20-14:40) Sensitive Detection of Trace Pesticide Residue Implemented by Terahertz All-Dielectric Metamaterial ··**433**

Zijian Cui   (Xi'an University of Technology)

Yue Wang (Xi'an University of Technology)

Xiaoju Zhang (Xi'an University of Technology)

Lisha Yue (Xi'an University of Technology)

Xinmei Wang (Xi'an University of Technology)

Chen Yang (Xi'an University of Technology)

Hui Hu (Xi'an University of Technology)

Kuang Zhang (Harbin Institute of Technology)

Yue Wang (Harbin University of Science and Technology)

(14:40-15:00) Super-Resolution Imaging Using Very Deep Convolutional Network in Terahertz NDT Field··**438**

Qiang Wang (Air Force Engineering University)

Hongbin Zhou (Air Force Engineering University)

Yi Wang (Air Force Engineering University)

Ruicong Xia (Air Force Engineering University)

Qiuhan Liu (Air Force Engineering University)

Boyan Zhao (Air Force Engineering University)

Yue Wang (Air Force Engineering University)

(14:00-14:20) Generative Model Driven Sampling Strategy for High Efficient Measurement of Complex Surfaces on Coordinate Measuring Machines ···442

Jieji Ren (Shanghai Jiaotong University)

Mingjun Ren (Shanghai Jiaotong University)

Lijian Sun (Zhejiang Lab)

(14:20-14:40) Architecture and Implementation of IoT Middleware for Ground Support Systems in Launch Site··448

Litian Xiao (Beijing Special Engineering Design and Research Institute)

Nan Xiao (China Aerospace Academy of System Science and Engineering)

Mengyuan Li (Beijing Special Engineering Design and Research Institute)

Shanshan Xie (Information Support Lab, Equipment Department, Aerospace System Ministry)

Kewen Hou (Beijing Special Engineering Design and Research Institute)

Yuliang Li (Beijing Special Engineering Design and Research Institute)

(14:40-15:00) Dual-lane Phononic Crystal for Low-frequency Elastic Wave Attenuation·········455

Jiawen Xu (Southeast University)

Ruqiang Yan (Xi'an Jiaotong University)

(15:00-15:20) Research Progress of Intelligent Lightning Protection Based on Internet of Things Technology ···459

Qibin Zhou(Shanghai University)

Feipeng Tang (Shanghai University)

Xin Huang (Shanghai University)

Huahui Chen(Shanghai Meteorological Bureau)

(15:20-15:40) Parameter Study of the Variable Reluctance Energy Harvester for Smart Railway Axle Box Bearing··464

Yun Gong (Chongqing university)

Sijia Wang (Chongqing university)

Zhengqiu Xie (Chongqing university)

Wenbin Huang (Chongqing university)

Poster Presentations

Diancong Zhang (Qufu Normal University)

Dianguo Cao (Qufu Normal University)

Jingyu Zhang (Qufu Normal University)

Jiashuai Wang (Qufu Normal University)

Jingxin Cao (Qufu Normal University)

Xi Chen (Qufu Normal University)

Ran Wei (Anhui University)

Yawei Hu (Anhui University)

Changbo He(Anhui University)

Zheng Cao (Anhui University)

Siliang Lu(Anhui University)

Fang Liu(Anhui University)

Yongbin Liu (Anhui University)

Zhichao Wu(Anhui University)

Jiahao Niu (Anhui University)

Siliang Lu (Southwest Jiaotong University)

Yongbin Liu (Anhui University)

Fang Liu (Anhui University)

Hao Sheng (Hunan University of Technology)

Zhongsheng Chen (Hunan University of Technology)

Yemei Xia (Hunan University of Technology)

Weimin Wang (Beijing University of Chemical Technology)

Liping Zhao (Xi'an Jiaotong University)

Xiangqian Cheng (Xi'an Jiaotong University)

Yiyong Yao (Xi'an Jiaotong University)

Xiaoming Wang (Beijing University of Chemical Technology)

Zhengcai Guo (Beijing University of Chemical Technology)

Guoan Yang (Beijing University of Chemical Technology)

Tianyu Gao (Harbin Institute of Technology)

Jingli Yang (Harbin Institute of Technology)

Shouda Jiang (Harbin Institute of Technology)

Jiawen Xu (Southeast University)

Xin Zhang (Southeast University)

Ruqiang Yan(Xi'an Jiaotong University)

Meiru Liu (Xi'an Jiaotong University)

Weiqiang Gao (AECC Sichuan Gas Turbine Establishment)

Xiao Xiao (AECC Sichuan Gas Turbine Establishment)

Guangrong Teng(AECC Sichuan Gas Turbine Establishment)

Chunyan Ao (Xi'an Jiaotong University)

Baijie Qiao (Xi'an Jiaotong University)

Yang Liu (Chinese Academy of Sciences)

Yuanyuan Dai (Chinese Academy of Sciences)

Guohua Liu (Harbin Institute of Technology)

Jingli Yang (Chinese Academy of Sciences)

Longfei Tian (Chinese Academy of Sciences)

Hua Li (Shanghai Jiaotong University)

Kuo Gao (Chinese Academy of Sciences)

Huijie Liu (Chinese Academy of Sciences)

Lei Liu (Chinese Academy of Sciences)

Xiaotong Pan (Chinese Academy of Sciences)

Hao Wang (Chinese Academy of Sciences)

Yang Liu (Chinese Academy of Sciences)

Zhiyong Li (CSG Power Dispatching and Control Center)

Chunxiao Liu (CSG Power Dispatching and Control Center)

Chi Tan (Xi'an Jiaotong University)

Yanzhen Zhao (Xi'an Jiaotong University)

Zijian Tang (Xi'an Jiaotong University)

Xuefeng Chen (Xi'an Jiaotong University)

Yafei Xu (Xi'an Jiaotong University)

Chao Ren (University of Electronic Science and Technology of China)

Lijian Zhu (Shanghai Space Propulsion Technology Research Institute)

Libing Bai (University of Electronic Science and Technology of China)

Lulu Tian (University of Electronic Science and Technology of China)

Jie Zhang (University of Electronic Science and Technology of China)

Zhen Liu (University of Electronic Science and Technology of China)

Jiangshan Ai (University of Electronic Science Technology of China)

Sihua Chen (Shanghai Space Propulsion Technology Research Institute)

Peng Deng (Sichuan Aerospace Measurement & Test Research Institute)

Libing Bai(University of Electronic Science Technology of China)

Lulu Tian (University of Electronic Science Technology of China)

Jie Zhang (University of Electronic Science Technology of China)

Yukun Zhang (Shaanxi Normal University)

Jisheng Li (Shaanxi Normal University)

Hanchao Liu (Shaanxi Normal University)

Rong Liu (Shaanxi Normal University)

Fan Yang (Shaanxi Normal University)

Ting Li (Shaanxi Normal University)

Zhigang Li (Southeast University)

Jianwei Cui (Southeast University)

Pudong Lu (Southeast University)

Huice Jiang   (Southeast University)

Guiping Lu (Beijing Institute of technology, Zhuhai)

Zhiyong Xiao (Beijing Institute of technology, Zhuhai)

Zhensheng Zhong (Beijing Institute of technology, Zhuhai)

Jiaran Liang (Beijing Institute of technology, Zhuhai)

Zhili Xiang (Beijing Institute of technology, Zhuhai)

Zhiyang Tang (Beijing Institute of technology, Zhuhai)

Jing Li (Southeast University)

Aidong Deng (Southeast University)
Yong Yang (Southeast University)
Jing Zhu (Southeast University)
Minqiang Deng (Southeast University)

Xiaojuan Zhao (State Grid Chongqing Electric Power Research Institute)
Yaqi Zhou(State Grid Hubei Transmission and Transformation Engineering CO.LTD)
Xue Wang (32620 Army)
Ruilin Xu (State Grid Chongqing Electric Power Research Institute)
Shihai Han (State Grid Chongqing Electric Power Research Institute)
Xin Chen (State Grid Hubei Transmission and Transformation Engineering CO.LTD)

Author Index

A Fuzzy Logic-based Method for Proton Exchange Membrane Fuel Cell Fault Diagnosis

Mengliu Pei
Department of Precision Machinery and Precision Instrumentation
University of Science and Technology of China
Hefei, China
mengliup@126.com

Chen Zhang
Department of Precision Machinery and Precision Instrumentation
University of Science and Technology of China
Hefei, China
zhang985@mail.ustc.edu.cn

Mingruo Hu
School of Mechanical Engineering
Shanghai Jiao Tong University
Shanghai, China
mingruohu@sjtu.edu.cn

Lisa Jackson
Department of Aeronautical and Automotive Engineering
Loughborough University
Loughborough, UK
l.m.jackson@lboro.ac.uk

Lei Mao*
Department of Precision Machinery and Precision Instrumentation
University of Science and Technology of China
Hefei, China
leimao82@ustc.edu.cn

Abstract—**Due to the fact that proton exchange membrane fuel cell system application is restricted by its limited reliability and durability, an efficient technique for accurate and fast identification of PEM fuel cell faults is highly required. In this study, a PEM fuel cell fault diagnostic method is proposed based on fuzzy logic model and electrochemical impedance spectroscopy. The fuzzy rules used in fuzzy logic model for inferring diagnosis are derived directly from PEM fuel cell faulty mechanisms, thus avoiding inconsistent and even conflicting fuzzy rules used in previous studies. Moreover, its effectiveness in identifying different PEM fuel cell faults and various levels of faults are investigated. Results demonstrate that PEM fuel cell flooding and dehydration can be diagnosed accurately and PEM fuel cell faults at different levels can also be discriminated.**

Keywords—Proton exchange membrane fuel cell, fault diagnosis, fuzzy logic, electrochemical impedance spectroscopy, water management

I. INTRODUCTION

Fuel cells are attracting more attention as they can offer environmentally friendly electrical power. Among different types of fuel cells, proton exchange membrane (PEM) fuel cell are considered to be promising due to its advantages such as high efficiency and no other emissions other than water. However, the reliability and durability of PEM fuel cell systems are still restricting the wider commercial applications [1]. Although significant progress has been made during recent years in PEM fuel cell technology, there still exist a huge gap with the target.

Because the PEM fuel cell is a complex multi-domain knowledge system and its performance can be greatly influenced by operating conditions, a variety of failures can occur during its operation, including poor water management issues, fuel starvation, etc. [2]. Therefore, it is necessary to detect and isolate PEM fuel cell faults, so as to take mitigation strategies for improving its reliability and durability.

In recent years, a set of researches have been devoted for PEM fuel cell diagnosis, and the widely used approaches include model-based methods [3] and data-driven methods [4]. In model-based methods, the mathematical model of PEM fuel cell should be developed, where in-depth knowledge of behavior inside PEM fuel cell is required. This restricts the use of model-based methods. With regard to data-driven method, signal processing techniques can be used to the test data for fault identification, without the requirement of PEM fuel cell mathematical model. However, sufficient test data at different PEM fuel cell status should be acquired, which is time-consuming and may not be possible in some situations, like obtaining sufficient test data from PEM fuel cell vehicles at faulty states. Moreover, as different PEM fuel cell systems may show performance variations due to system dimension and material property difference, the same fault will cause different levels of degradations at different PEM fuel cell systems, thus data-driven approaches might not be robust in PEM fuel cell fault diagnosis.

Besides the above two types of methods, knowledge-based approaches have also been applied for PEM fuel cell fault diagnosis, such as fuzzy inference system (FIS) which based on fuzzy logic (FL) and fuzzy set theory. It can imitate human reasoning to take decisions, and has been used in different fields including system state inference, process simulation and diagnosis [5]. In FL, inputs and corresponding fuzzy rules which in the form of IF-THEN statements are the key for accurate fault diagnosis. From previous studies, the FL model inputs can be generally divided into two groups, including features from characterization test like polarization curve or electrochemical impedance spectroscopy (EIS), and data collected during PEM fuel cell operation, such as current, voltage, temperature, etc. [6]. Among these inputs, features from EIS are more widely used, since these variables can be linked to PEM fuel cell internal mechanisms, thus the faults and associated degradations can be connected.

However, in most studies, the selection of EIS features as FL model inputs is still largely depend on researchers' experience. The reason is that in these studies, fuzzy rules associated with FL model inputs are generated using algorithms such as fuzzy clustering [7]. Therefore, the generated fuzzy rules are heavily rely on the training data and may not have physical meaning, which hinders the direct link between fuzzy rules and PEM fuel cell faults and associated degradation mechanisms. Moreover, with different EIS features selected, diagnostic performance bias may be caused, which brings great difficulty in applying knowledge-based approached in fault diagnosis at practical PEM fuel cell applications. Therefore, the appropriate selection FL model inputs for consistent diagnostic performance at PEM fuel cell systems is highly required.

In this study, a FL model for PEM fuel cell fault diagnosis is presented, where a novel method for generating fuzzy rules is proposed. In the analysis, the relationship between EIS features and PEM fault mechanisms is clarified, from which

---

*Lei Mao is the corresponding author. (e-mail: leimao82@ustc.edu.cn).

EIS features more affected by PEM fuel cell faults are selected as FL model inputs. With selected model inputs and clarified relationship, consistent fuzzy rules can be generated for the inference. Moreover, the PEM fuel cell water management issues, including flooding and dehydration, are selected in the study to investigate effectiveness of proposed method. The reason why choose water management issues is that these issues are usually experienced in practical PEM fuel cell applications [8], and with proper mitigation strategies, the degradation caused by poor water management can be recovered efficiently. Test data is collected such that the robustness of proposed method can be better highlighted.

The contribution of this work lies in that FL model inputs are selected from expert knowledge or clarified results in previous studies, which can directly represent the PEM fuel cell performance variation. Therefore, the generated fuzzy rules using these inputs can be linked to PEM fuel cell fault mechanisms directly. Consequently, its performance and robustness can be ensured, which will be beneficial in fault diagnosis at practical PEM fuel cell applications. The flowchart of developed FL model is shown in Fig. 1.

| Clarification of EIS features | ⇨ | Selection of FL model inputs | ⇨ | Generation of Fuzzy rules | ⇨ | Construction of FL model |
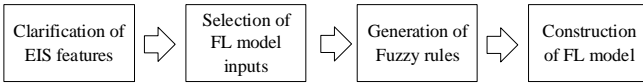
Fig. 1. Illustration of proposed method for PEM fuel cell fault diagnosis.

This paper is organized as follows. In section II, the features extracted from EIS spectrum are selected based on previous studies, and development of fuzzy rules and FL model are also presented. Model validation, including the description of the test data under different faults with different levels and the diagnosis results of the proposed method is given in section III. From the findings, conclusions are made in section IV.

II. METHODOLOGY

A. Selection of EIS Features

EIS is a versatile tool for fuel cell diagnosis that can provide rich information for evaluating fuel cell performance [6]. In the analysis, features from EIS more affected by PEM fuel cell faults will be extracted and used as FL model inputs. This can be achieved from previous studies regarding PEM fuel cell degradation mechanisms. As described in section I, PEM fuel cell water management issues are investigated in this study, thus EIS features affected by flooding and dehydration are determined, with which fuzzy rules are generated for fault diagnosis.

A typical EIS spectrum contains a high frequency loop and a low frequency loop. The high frequency loop corresponding to the charge transfer process of PEM fuel cell and the low

frequency loop is in relation with the mass transport process of PEM fuel cell [9].

Based on previous studies, several EIS features sensitive to PEM fuel cell poor water management issues are determined, which are summarized as follows.

Internal resistance (R_m). The internal resistance represents the total Ohmic resistance of PEM fuel cell, among which is mainly the membrane resistance, which can reflect the membrane humidification level [4]. In the General Motors patent [10], the interdependency between the degree of humidification and the internal resistance of the fuel cell was studied, where proposed that the internal resistance would be increased with cell drying. While in [11], it was concluded that the internal resistance was only slightly modified with the PEM fuel cell flooding.

Charge transfer resistance (R_t). Many researchers connected the diameter of the high frequency loop of EIS with the charge transfer resistance [12]. In [13], Hsieh et al. found that the charge transfer resistance decreased due to the increased dehydration. This is also in agreement with [14], where it also found that water accumulation not only impeded the transfer of electron at the electrode surface but also reduced the electron transfer rate, leading an increase of the transfer resistance.

Warburg impedance (W). The resistance describes the diffusion convection of the different gases at the electrode, which can be represented by diffusion Ohmic resistance (W-r) and diffusion time constant (W-τ). In [15], the authors pointed out the diffusion Ohmic resistance increased with the increase of water content, while the decrease of diffusion time constant was also observed.

Double layer capacitance (C). The capacitance is at interface between electrode and electrolyte, which will increase with the water accumulation according to [15].

Based on above studies, EIS features changed in the process of dehydration and flooding are summarized in the Table I, where *'I', 'D'* represent increase and decrease, respectively, and *'C'* represents constant or slightly changed during the flooding or membrane dehydration.

B. Generation of Fuzzy Rules

FL starts with the concept of fuzzy set, which has not a clearly defined boundary and contains only a partial degree of membership compared with the classical set. The overlap between these sets indicates how behavior changes under different operating conditions, and the centers of these sets can be defined by known behavior from the rules-literature [16]. In this study, input fuzzy sets are defined as the percentage changes of EIS features associated with PEM fuel cell faults, where a positive value indicates that faults can cause feature

TABLE I.   EIS FEATURE ARIATIONS UNDER FLOODING AND DEHYDRATION SITUATION

	Flooding	Dehydration
Internal resistance (R_m)	*D* or *C*	*I*
Charge transfer resistance (R_{ta} and R_{tc})	*I*	*D*
Diffusion Ohmic resistance (W-r)	*I*	*D*
Diffusion time constant (W-τ)	*D*	*I*
Double layer capacitance (C_{da} and C_{dc})	*D*	*I* or *C*

increase, and vice versa. It should be noted that, for simplification purpose, triangular and trapezoidal set functions are used as membership functions herein. The fuzzy input sets of the FL model are shown in Fig. 2.

As for the output fuzzy set, a scale is developed to indicate different PEM fuel cell status, i.e., normal situation, flooding and membrane dehydration, which is shown in Fig. 3. In this study, a 0 - 1 range is proposed with a rating of between 0.5 and 0.6 representing the "normal" situation, lower numbers below 0.5 indicating a 'dehydration' failure, and scale numbers higher than 0.6 representing a 'flooding' failure. Similarly, triangular and trapezoidal set functions are used.

Based on the fuzzy sets and selected EIS features, fuzzy rules can be generated with the form of logical IF-THEN statements and are listed in Table II. These rules are derived directly from previous studies, thus have physical meanings of degradation mechanisms. Moreover, although only fuzzy rules related to poor PEM fuel cell water managements are generated herein, they can be easily expanded to include more PEM fuel cell faults.

C. Framework of Proposed Diagnostic Technique

From above results, FL model inputs can be determined with selected EIS features. Based on the fuzzy rules which extracted from previous studies directly, inference can be made to distinguish the PEM fuel cell states. The framework of the diagnostic technique is depicted in Fig. 4.

III. VALIDATION

For investigating the effectiveness and robustness of proposed method, two groups of EIS data are used from [17], including EIS data collected at different PEM fuel cell faults (flooding and dehydration herein), and EIS data collected at various levels of faults. With the test data, diagnostic capacity of proposed method in discriminating both different PEM fuel cell faults and various fault levels can be clarified.

In this section, the test data is firstly presented. In order to extract EIS features described in section II, an equivalent circuit model (ECM) is used to fit these data. With extracted EIS features, the diagnostic performance of proposed method is then investigated.

A. Acquisition of the Test Data

In this study, test data from previous study [17] is used, which can be regarded as the benchmark data for validating the effectiveness of proposed method. In the test, a six-cell PEM fuel cell stack with an active surface area of 150 cm^2 was adopted, both flooding and membrane dehydration were triggered on-line. For the EIS measurements, ac perturbation was chosen as a function of the direct current. In addition, the frequencies were ranged from 0.1 Hz to 1 kHz, with 10 measurement points per decade were applied.

The first group of data, i.e., the data acquired under three different status of PEM fuel cell is recorded at 70 A dc, as shown in Fig. 5.

The second group of data, is shown in Fig. 6 (a) and (b). These EIS spectrum measurements were taken during the experiments where the mean cell voltage evolution was recorded while the water built up or the membrane got drying and recorded as minor, medium, and serious according to the severity of degradation. The arrangements of the system in

the experiment with water built up is listed in Table III, where h represents the inlet gases relative humidity, T represents the temperature and S represents the stoichiometry imposed at cathode or anode. For trigging the membrane dehydration, the inlet gas relative humidity lowered to 10% and 15% on anode and cathode side, respectively, with other conditions remain the same with the flooding one. EIS spectrum was recorded at 70 A for both flooding and membrane dehydration.
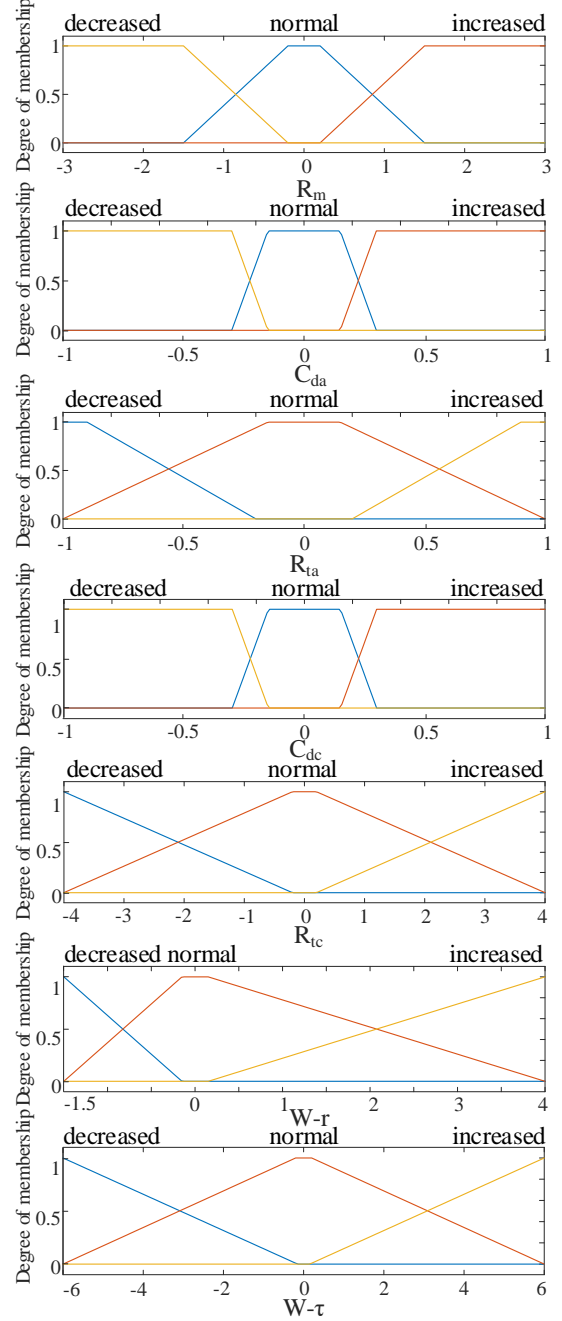
Fig. 2.   Fuzzy input sets with membership functions.
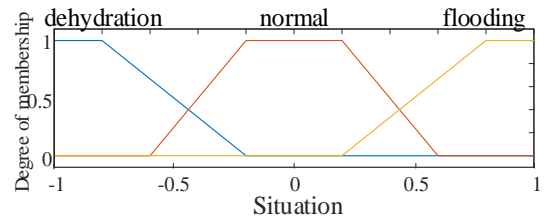
Fig. 3.   Fuzzy output set with membership function.

3

TABLE II.   GENERATED FUZZY RULES

	IF	THEN
1	R_m is nomal AND C_{da} is normal AND R_{ta} is nomal AND C_{dc} is normal AND R_{tc} is normal AND W-r is normal AND W-τ is normal	Situation is normal
2	R_m is increased AND C_{da} is normal AND R_{ta} is nomal AND C_{dc} is normal AND R_{tc} is normal AND W-r is normal AND W-τ is normal	Situation is dehydration
3	R_m is increased OR C_{da} is decreased OR R_{ta} is decreased OR C_{dc} is decreased OR R_{tc} is decreased OR W-r is decreased OR W-τ is increased	Situation is dehydration
4	R_m is not increased OR C_{da} is not decreased OR R_{ta} is increased OR C_{dc} is not decreased OR R_{tc} is increased OR W-r is increased OR W-τ is decreased	Situation is flooding
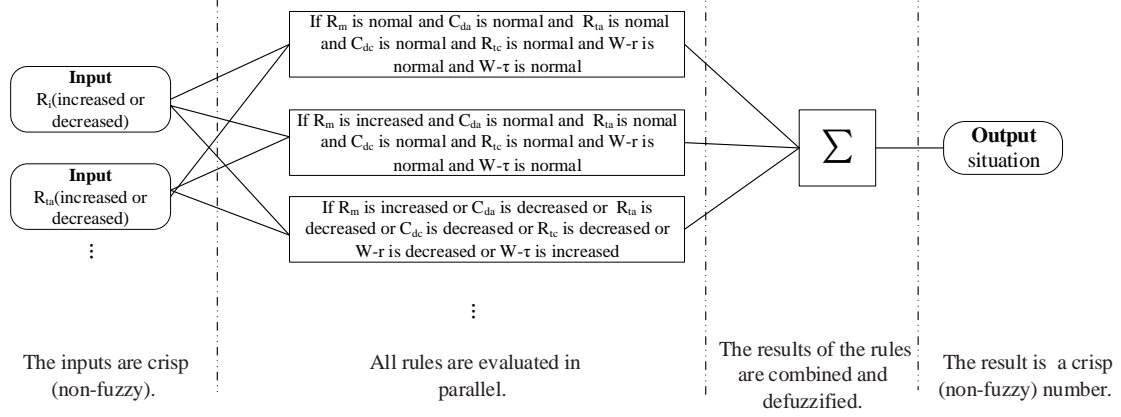
Fig. 4. Fuzzy diagnostic system structure.
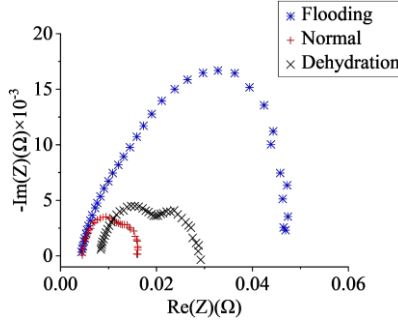
Fig. 5. EIS spectrum recorded at 70A under three situations.

B. Equivalent Electrical Model of the PEM Fuel Cell

In this paper, the ECM containing features extracted in section II is shown in Fig. 7, which is selected to fit the test data. It retains a strong physical meaning and is found to fit very well to the data. It has been widely used in previous studies and its effectiveness have been proved in [18]. For each spectrum, the model is fitted to the data using ZView software in the analysis.

C. Capacity to Distinguish Different Fault

To test the efficiency and robustness of the FL model in distinguishing the different situations of PEM fuel cell, the first group data (i.e. Fig. 5) is fitted to the ECM, and Table IV lists the variations of extracted features at each PEM fuel cell status and △ represents the changes of features in faulty status compared with normal status. As shown in Table IV, the selected EIS features are clearly affected with PEM fuel cell flooding and membrane dehydration, indicating that the selected FL model inputs are sensitive to PEM fuel cell faults.

The variations of extracted EIS features are then used as inputs to the constructed FL model (as depicted in Fig. 4),
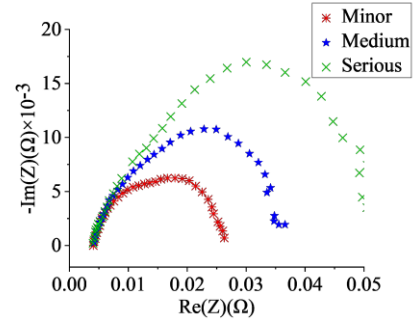
(a) EIS spectrum while the water built up.
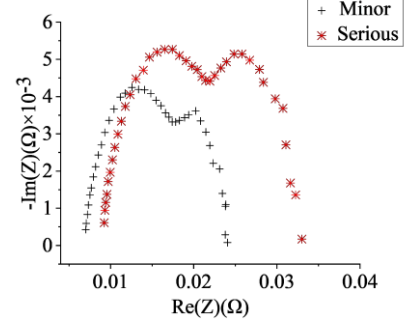
(b) EIS spectrum while the membrane got drying.

Fig. 6. EIS spectrum under different degree of faulty situaions.

TABLE III.   THE ARRANGEMENT OF THE EXPERIMENT WITH WATER BUILT UP

h_{ca} (%)	h_a (%)	T (°C)	S_{ca}	S_a
50	70	60	4	1.2

where fuzzy rules listed in Table II are incorporated, from which the PEM fuel cell states can be inferred. The outputs of the FL model in different situations are listed in Table V.

From Table V, both flooding and dehydration can be discriminated accurately with the proposed method. Since no training process and manual configuration are required in the method, the proposed technique can be used as fast and automatic fault identification tool in practical PEM fuel cell systems.
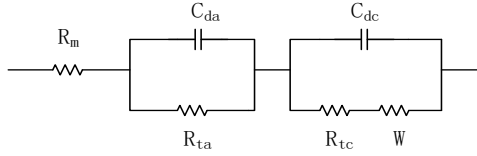
D. Capacity to Distinguish Different Degree of the Fault

To test the capability of the model in discriminating different degrees of flooding or dehydration, the second group of data (as shown in Fig. 6) is fitted to the ECM. Similarly, the features and corresponding increments presented by △ while water built up and the membrane got drying are given in Table VI and Table VII, respectively. It can be seen from above tables that different levels of PEM fuel cell faults, especially the minor fault, can also significantly affect the selected EIS features, indicating the potential of proposed method in identifying early stage PEM fuel cell faults.

Fig. 7. ECM of PEM fuel cell.

TABLE IV. THE FEATRUES AND CORRESPONDING VARIATIONS IN FAULTY STATUS COMPARED WITH NORMAL STATUS

	R_m	C_{da}	R_{ta}	C_{dc}	R_{tc}	W-r	W-τ
Normal state	0.0045	6.371	0.0094	0.4337	0.0056	0.0035	0.0808
Flooding	0.0047	6.9490	0.0320	0.5332	0.0090	0.0085	0.0700
$\triangle_1$(%)	4.4	9.1	240.4	22.9	60.7	142.9	-13.4
Dehydration	0.0149	1.3760	0.0057	0.1971	0.0024	0.0029	0.1552
$\triangle_2$ (%)	231.1	-78.4	-39.4	-54.6	-57.1	-17.1	92.0

TABLE V. DIAGNOSTIC RESULTS WITH CONSTRUCTED FL MODEL IN THREE DIFFERENT STATUS

	Normal	Flooding	Dehydration
The output of the FL	0.5	0.863	0.137

TABLE VI. EVOLUTION OF THE FEATURES WHILE WATER BUILDS UP

	R_m	C_{da}	R_{ta}	C_{dc}	R_{tc}	W-r	W-τ
Minor	0.0046	6.9150	0.0103	0.3901	0.0130	0.0072	0.0536
$\triangle_2$ (%)	1.4	8.5	9.0	-10.1	131.7	106.7	-33.7
Medium	0.0046	7.3200	0.0129	0.3765	0.0210	0.0146	0.0509
$\triangle_3$ (%)	3.2	14.9	37.1	-13.2	274.9	317.9	-37.0
Serious	0.0046	7.4210	0.0153	0.4752	0.0324	0.0155	0.0365
$\triangle_4$ (%)	2.3	16.5	63.1	9.6	479.3	341.5	-54.8

TABLE VII. EVOLUTION OF THE FEATURES THE MEMBRANE GETTING DRY

	R_m	C_{da}	R_{ta}	C_{dc}	R_{tc}	W-r	W-τ
Minor	0.0084	5.0900	0.0073	0.3205	0.0055	0.0025	0.1429
$\triangle_2$ (%)	87.5	-26.1	-22.8	-20.1	-1.4	-27.9	76.8
Serious	0.0096	4.0256	0.0065	0.3033	0.0043	0.0011	0.1244
$\triangle_3$ (%)	114.0	-30.1	-31.3	-36.8	-22.9	-68.9	53.9

TABLE VIII. DIAGNOSTIC RESULTS WITH WATER BUILDS UP

	Minor	Medium	Serious
The output of the FL	0.656	0.716	0.863

TABLE IX. DIAGNOSTIC RESULTS THE MEMBRANE GET DRYING

	Minor	Serious
The output of the FL	0.487	0.331

With extracted EIS features and fault diagnosis procedure depicted in Fig. 4, outputs of the FL model for different degree of flooding and membrane dehydration are listed in Table VIII and Table IX, respectively.

It can be seen that PEM faults at different levels can also be distinguished with constructed FL model. This will be beneficial in practical applications, since with accurate identification of faults at different levels, especially the early stage fault, appropriate mitigations can be taken to recover PEM fuel cell performance before irreversible degradation is caused.

It can be concluded from above results that the proposed FL-based fault diagnostic technique has several advantages over the previous FL-based methods. 1) Training data is not required in the proposed method for generating fuzzy rules, this makes the proposed method be used automatically. 2) The fuzzy rules in the study are generated with clarified degradation mechanisms and expert knowledge, thus with consistent fuzzy rules, the proposed method can be used at different PEM fuel cell systems for accurate diagnosis. 3) Although only flooding and dehydration are investigated in this study, the proposed method can be easily expanded to accommodate more faults, by generating related fuzzy rules.

IV. CONCLUSIONS

In this paper, a PEM fuel cell fault diagnostic method is proposed based on FL model, and its effectiveness in identifying PEM fuel cell poor water management issues is investigated.

In the analysis, the fuzzy rules are derived directly from PEM fuel cell faulty mechanisms, where relationships between PEM fuel cell faults (flooding and dehydration herein) and affected PEM fuel cell features in EIS are obtained. The effectiveness of proposed method is investigated using PEM fuel cell test data at different status and in different degrees. Results demonstrate that not only PEM fuel cell flooding and dehydration can be discriminated, but also PEM fuel cell faults in different degrees can be identified with the proposed method.

The proposed method can be great beneficial in practical applications. Firstly, since training data is not used in generating fuzzy rules, whose obtaining is time-consuming and expensive, the proposed method can be used more efficiently in fault diagnosis. Secondly, its capability of identifying early stage PEM fuel cell fault can be used to extend the system lifetime, as mitigation strategies can be taken to remove the fault effect. Last but not the least, the proposed method is expandable, indicating the fuzzy rules relating to more PEM fuel cell faults can be generated at the same manner, such that the proposed method can be used to identify more faults.

REFERENCES

[1] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Prognostics and Health Management of PEMFC–State of the art and remaining challenges," Int. J. Hydrogen Energy, vol. 38, no. 35, pp. 15307-15317, 2013.

[2] R. Borup et al., "Scientific Aspects of Polymer Electrolyte Fuel Cell Durability and Degradation," Chem. Rev., vol. 38, no. 50, pp. no-no, 2010.

[3] A. Hernandez, D. Hissel, and R. Outbib, "Modeling and fault diagnosis of a polymer electrolyte fuel cell using electrical equivalent analysis," IEEE Transactions on Energy Conversion, vol. 25, no. 1, pp. 148-160, 2009.

[4] Z. Zheng, M.-C. Péra, D. Hissel, M. Becherif, K.-S. Agbli, and Y. Li, "A double-fuzzy diagnostic methodology dedicated to online fault diagnosis of proton exchange membrane fuel cell stacks," J. Power Sources, vol. 271, pp. 570-581, 2014.

[5] T. J. Ross, Fuzzy logic with engineering applications. John Wiley & Sons, 2005.

[6] N. Yousfi-Steiner, P. Moçotéguy, D. Candusso, D. Hissel, A. Hernandez, and A. Aslanides, "A review on PEM voltage degradation associated with water management: Impacts, influent factors and characterization," J. Power Sources, vol. 183, no. 1, pp. 260-274, 2008.

[7] Y. Liu, S. Yi, and Z. Liu, "An approach to fault diagnosis for non-linear system based on fuzzy cluster analysis," in IEEE Instrumentation & Measurement Technology Conference, 2000.

[8] M. Ji and Z. Wei, "A review of water management in polymer electrolyte membrane fuel cells," Energies, vol. 2, no. 4, pp. 1057-1106, 2009.

[9] Perry and L. Mike, "Mass Transport in Gas-Diffusion Electrodes: A Diagnostic Tool for Fuel-Cell Cathodes," J. Electrochem. Soc., vol. 145, no. 1, p. 5, 1998.

[10] M. F. Mathias and S. A. Grot, "System and method for controlling the humidity level of a fuel cell," ed: Google Patents, 2002.

[11] S. Rodat, S. Sailler, F. Druart, P.-X. Thivel, Y. Bultel, and P. Ozil, "EIS measurements in the diagnosis of the environment within a PEMFC stack," J. Appl. Electrochem., vol. 40, no. 5, pp. 911-920, 2010.

[12] X. Yuan, C. S. Jian, H. Wang, and J. Zhang, "AC impedance diagnosis of a 500 W PEM fuel cell stack. Part II: Individual cell impedance," J. Power Sources, vol. 161, no. 2, pp. p.929-937, 2006.

[13] S.-S. Hsieh, S.-H. Yang, and C.-L. Feng, "Characterization of the operational parameters of a H2/air micro PEMFC with different flow fields by impedance spectroscopy," J. Power Sources, vol. 162, no. 1, pp. 262-270, 2006.

[14] J.-D. Kim, Y.-I. Park, K. Kobayashi, M. Nagai, and M. Kunimatsu, "Characterization of CO tolerance of PEMFC by ac impedance spectroscopy," Solid State Ionics, vol. 140, no. 3-4, pp. 313-325, 2001.

[15] S. Laribi, K. Mammar, Y. Sahli, and K. Koussa, "Analysis and diagnosis of PEM fuel cell failure modes (flooding & drying) across the physical parameters of electrochemical impedance model: Using neural networks method," Sustain. Energy. Techn., vol. 34, pp. 35-42, 2019.

[16] B. Davies, L. Jackson, and S. Dunnett, "Expert diagnosis of polymer electrolyte fuel cells," Int. J. Hydrogen Energy, vol. 42, no. 16, pp. 11724-11734, 2017.

[17] N. Fouquet, C. Doulet, C. Nouillant, G. Dauphin-Tanguy, and B. Ould-Bouamama, "Model based PEM fuel cell state-of-health monitoring via ac impedance measurements," J. Power Sources, vol. 159, no. 2, pp. 905-913, 2006.

[18] S. Rodat, S. Sailler, F. Druart, P. X. Thivel, Y. Bultel, and P. Ozil, "EIS measurements in the diagnosis of the environment within a PEMFC stack," J. Appl. Electrochem., vol. 40, no. 5, pp. 911-920, 2010.

On-line Fault Diagnosis of Proton Exchange Membrane Membrane Fuel Cell Using Polarization Curve

Shuyou Wu
Wuhan Second Ship Design and Research Institute
Wuhan, China
39752110@qq.com

Xu Peng
Wuhan Second Ship Design and Research Institute
Wuhan, China
18907159492@189.cn

Lei Mao*
Department of Precision Machinery and Precious Instrumentation
University of Science and Technology of China
Hefei, China
leimao82@ustc.edu.cn

Ao Liu
Wuhan Second Ship Design and Research Institute
Wuhan, China
liuaolow@mail.ustc.edu.cn

Liang Dong
Wuhan Second Ship Design and Research Institute
Wuhan, China
472125186@qq.com

Tong Wang
Wuhan Second Ship Design and Research Institute
Wuhan, China
674458148@qq.com

Abstract—**The study investigates performance of using polarization curve for fuel cell fault detection and isolation. As polarization curve can be obtained at practical proton exchange membrane fuel cell (PEMFC) system efficiently, its diagnostic procedure can be used for on-line purpose. The paper starts with background knowledge about polarization curve and its use in fault diagnosis. Polarization curves obtained at a PEMFC system will then be employed to detect abnormal behavior. By studying variation in the gradients of collected curves, PEMFC fault is identified. Furthermore, two classification techniques are applied to isolate different PEMFC conditions, including healthy and faulty conditions. Results demonstrate that by monitoring and analyzing polarization curves, PEMFC fault can be identified accurately, and processing of such information can meet on-line diagnosis requirements.**

Keywords—PEMFC, fault diagnostics, polarization curve, wavelet packet transform

I. INTRODUCTION

Fuel cell is the core component of new energy vehicle, which is related to the vehicle's endurance and other key performance. With development of hydrogen and fuel cell technology, they have been equipped at different applications. However, PEMFC operating safety and its useful lifetime are still major barriers preventing it from further commercialization.

In the last few decades, a set of studies have been focused on PEMFC fault identification, which can be separated into model based and non-model driven methods [1-11]. Regarding the model driven methods, PEMFC model should be constructed, system faults can then be detected and isolated with comparison of residual between model outputs and actual measurements. Three kinds of models can be constructed for this purpose. White box model uses space differential equations to express activities inside fuel cells, thus they can provide accurate diagnostic results. But due to their complexity and time-consuming analysis, they are not suitable for on-line diagnosis [3]. Grey box model is developed using physical laws derived from priori knowledge, thus mathematical formulas is replaced using empirical equations, which makes them commonly used in the studies to perform

diagnosis of fuel cell systems [4-6]. While black box model relies on data-based approaches, with training data, the relationship between inputs and outputs of the models are deduced. This developed relationship is then employed to perform fault diagnosis. Therefore, black-box models are usually used in complex non-linear systems, where it is difficult for model development using mathematical or empirical equations [7-8]. Besides model based techniques, non-model driven techniques is also utilized to identify faults in fuel cell systems. Compared to model based methods, non-model based techniques only rely on expert knowledge or test data. Based on previous studies, artificial intelligence techniques, like neural network [9], fuzzy logic [10], statistical methods [7-8], including principle component analysis, fisher discriminant analysis, kernel fisher discriminant analysis, and kernel principle component analysis, moreover, signal processing methods based on wavelet transform [11] and Fourier transform (FT), can be used to identify PEMFC abnormal conditions.

For applying above techniques for fault identification, PEMFC system sensor measurements should be collected and used for analysis, collected PEMFC data should be capable of representing behavior inside fuel cells, especially changes due to PEMFC faults. With previous researches, a series of sensors can be utilized in PEMFC diagnostics, including voltage, current, temperature, pressure, gas flow, etc. [7-10]

With PEMFC sensor measurements, PEMFC abnormal conditions can be identified using fault diagnostic algorithms proposed in previous studies. However, as large amount of data samples are obtained with PEMFC operation, fault diagnostic algorithms may be time-consuming, thus these proposed methodologies can only be used for off-line diagnosis. As a possible solution, several researches have been devoted on-line PEMFC diagnostics using single measurement, such as PEMFC output voltage [10] and electrochemical impedance spectroscopy (EIS) [11-12]. However, the use of voltage alone in diagnostics is only valid in steady-state conditions, as current effect must be considered in the dynamic conditions, while measurement of EIS requires extra testing equipment and may interrupt normal fuel cell operation, thus their applications for on-line fault diagnosis are constrained.

---

* Lei Mao is the corresponding author. (e-mail: leimao82@ustc.edu.cn).

According to previous studies, polarization curve can be collected efficiently from PEM fuel cell, and several losses, such as Mass transport loss, activation loss, and Ohmic loss, can be characterized with the curve. Compared to EIS, polarization curve can be collected without extra testing equipment, and normal operation of fuel cell system will not be interrupted [13-14]. However, further study is still required for investigating diagnostic effectiveness using polarization curve, especially at practical PEM fuel cell systems.

This study investigates effectiveness of using polarization curve in identifying practical PEMFC system faults. In section 2, background knowledge of polarization curves is described. Section 3 collects polarization curve at a practical PEMFC system, and gradients of activation and Ohmic losses from polarization curves are used to identify PEM fuel cell abnormal performance. Furthermore, two classification algorithms are applied to discriminate different fuel cell conditions. Section 4 provides the conclusions..

II. DESCRIPTION OF POLARIZATION CURVE FOR FAULT DIAGNOSIS

From previous studies, polarization curve can express the relation between PEMFC current and corresponding output voltage, which could be measured efficiently at PEMFC system without increasing measurement complexity [24, 25]. In this process, different levels of current densities should be applied, and corresponding fuel cell voltages will be collected.

It is noted that at polarization curve measurement process, for each current level, sufficient time (usually 5-15 minutes) is required for the system to reach stabilization. Moreover, as multiple current density and voltage points are collected, in order to get reliable polarization curves, the last 1/3 points for each current level should be averaged and then plotted to express the relationship between current density and voltage. The reason of averaging the last 1/3 points is that when the current jumps to a new level, certain time is required for the system stabilization, and the average operation can further improve the accuracy of polarization curve. Figure 1 depicts a typical PEMFC polarization curve.

It is found that three phases are included in the curve, which correspond to various PEMFC losses. With these losses, output voltage is calculated using the following equation.

$$V_{cell} = E_n - V_{act} - V_{ohm} - V_{mass} \qquad (1)$$

where V_{cell} is PEMFC output voltage (V), E_n is open circuit voltage , V_{act}, V_{ohm}, and V_{mass} represent the voltage losses due to different mechanisms, which are expressed as follows.
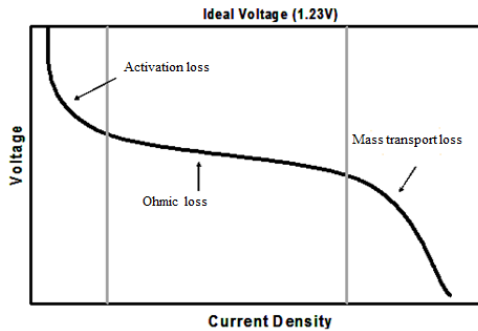
Fig. 1. Polarization curve of a PEM fuel cell

$$V_{act} = R\,T/(2{\times}F)\,ln(i/i_{oc}) \qquad (2)$$

$$V_{ohm} = i{\times}R_{membrane} \qquad (3)$$

$$V_{mass} = m_{mass}{\times}\exp(n_{mass}{\times}i) \qquad (4)$$

where i is current density (A/cm^2), R is universal gas constant (J/mol.K), T is temperature of PEMFC system (K), α is charge transfer coefficient, F is Faraday constant (C/mol), i_{oc} is the exchange current density at cathode (A/cm^2) ,$R_{membrane}$ is membrane resistance (Ω/cm^2), m_{mass} and n_{mass} are mass transport loss voltage coefficients.

From above results, it is expected that with PEMFC performance change, polarization curve is changed accordingly. Therefore, by analyzing different polarization curves, PEMFC condition could be evaluated effectively.

In the analysis, in order to minimize the computational time, two approaches are applied to polarization curves, including the direct comparison of polarization curve gradients, and extract feature from polarization curve using wavelet packet transform, which will be described below.

The first approach uses gradients of different phases from polarization curve to detect the PEMFC abnormal condition. It is expected that with existence of fuel cell fault, the corresponding phase of polarization curve would be affected. With this method, three gradients will be obtained from each polarization curve, and fuel cell fault is detected by comparing these gradients of different polarization curves.

In the second approach, wavelet packet transform (WPT) is selected to analyze polarization curve. With extracted wavelet coefficients, normalized energy is computed as follows.

$$E^p = \frac{1}{N_p} \sum_{j,k} \left| C_{j,k}^p \right|^2 \qquad (5)$$

where E^p is normalized energy of wavelet packet p, N_p is coefficient number in wavelet packet p, $C_{j,k}^p$ is WPT coefficient.

In this study, three layer wavelet packet transform is used, which is shown in Figure 2. It can be seen that 14 wavelet packets can be obtained from each polarization curve, from which 14 normalized energies are obtained and utilized for fault detection.
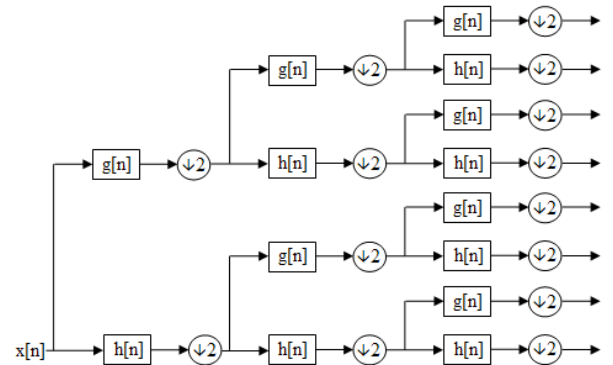
Fig. 2. 3 level WPT, g[n] and h[n] are low-pass and high-pass wavelet packets

It is mentioned that reason of selecting WPT herein is that with FT analysis, it is found that power spectrum of polarization curve is concentrated at low frequency range, thus features from low frequency domain should be used to capture PEMFC performance variation. With wavelet packet transform, coefficients are extracted from different frequency bins, and using normalized energies, information from various frequency bins can be expressed. This will be facilitate the extraction of more representative features for PEMFC diagnostics.

III. EFFECTIVENESS OF POLARIZATION CURVE IN PEMFC DIAGNOSTICS

In the analysis, an evaporative cooled (EC) PEMFC system is selected, which is depicted in Figure 3. It can be observed that during PEMFC operation, besides hydrogen and air, water is also input into PEMFC and partly evaporated to form a liquid/vapor mix carrying thermal energy. For maintaining PEMFC water balance, sufficient water is recovered in heat exchange and then injected to the water tank.
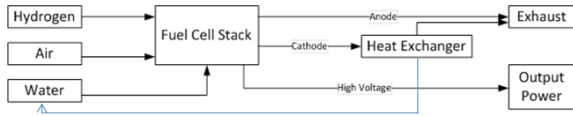
Fig. 3. Block diagram of EC fuel cell system

During PEMFC operation, the polarization curve is collected at certain time intervals, this is utilized for monitoring PEMFC performance. Figure 4 depicts a measured PEMFC polarization curve, with the collection procedure described in section 2.
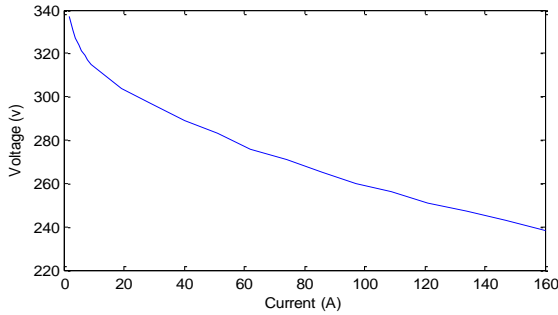
Fig. 4. A collected PEMFC polarization curve

In Figure 4, the third phase which expresses mass transport loss cannot be found. The reason is that in EC fuel cell system, the limit for maximum current is set to avoid the possible damage to electric elements in the system. Therefore, in this study, only activation loss and Ohmic loss are analyzed for PEMFC diagnostics.

As described before, behavior change in PEMFC system could be indicated using variations in polarization curve at different phases. Therefore, by monitoring evolutions of these phase gradients, the performance variation in PEM fuel cell system can be identified.

Figure 5 depicts evolutions of two gradients from collected polarization curves, including gradient of activation loss and gradient of Ohmic loss, where vertical red dashed lines represent the time point when PEMFC is replaced. At these time points, gradient evolution can be separated into four stages. It is mentioned that since PEMFC system was tested within the lab, thus the variation of external parameters, like environmental temperature and humidity, can be ignored in the analysis, and the variation of collected polarization curve is only due to the fuel cell performance change.
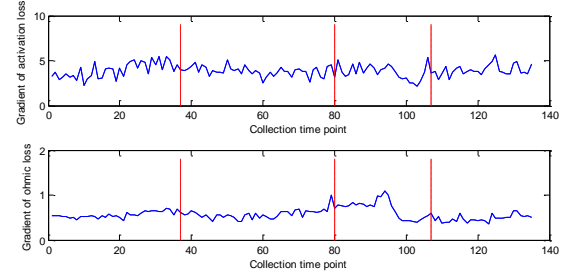
Fig. 5. Evolution of gradients for activation and Ohmic loss phases

Several findings can be observed from above figure. The gradients of Ohmic loss phase shows significant variation at stage 3 which is due to the hydrogen issue, while the evolution curve of activation loss phase is more 'noisy', and clearly reduction at stage 3 is also observed. It is mentioned that PEMFC stack replacement reduces the Ohmic loss and activation loss gradients slightly (shown in Figure 4 from stage 1 to stage 2), indicating the replaced fuel cell stack can still guarantee the fuel cell performance, this further confirm the necessity of using more effective health monitoring techniques in PEMFC systems.

As presented before, two approaches will be applied to perform fault detection using collected polarization curves. In the first approach, two gradients from each polarization curve are obtained, these values will be used to detect the existence of fuel cell fault. Figure 6 depicts the results of fault detection using two gradients. It can be seen that by comparing polarization curve gradient directly, PEMFC fault is not accurately identified, and several PEM fuel cell conditions are misclassified.
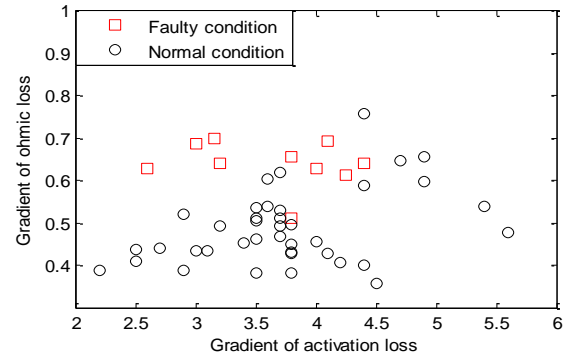
Fig. 6. Fault detection results using gradients from polarization curves

WPT is also utilized to analyze measured curves for extracting WP coefficients, and normalized energy is obtained. Figure 7 shows the fault detection results using two features (normalized energies), which are selected using SVD described in section 2. It should be noted that in the analysis, among 14 calculated energies, two highest values are used in diagnostics. Therefore, features 1 and 2 in Figure 7 corresponds to two highest energy value calculated using Eq. (5).

From Figure 7, by using two highest energy values, most fuel
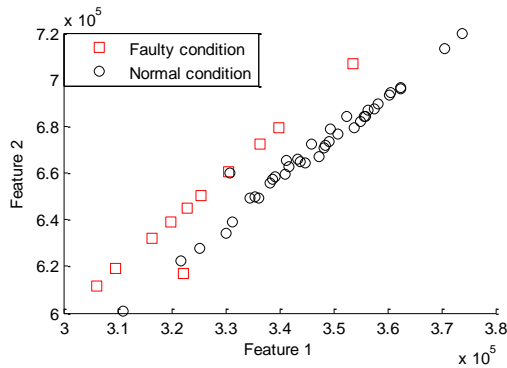
Fig. 7. Fault detection results using normalized energies from WPT

cell fault can be detected, as this detection can be made quickly (about 20 seconds), it could be used at practical PEMFC systems effectively to alert existence of the fuel cell fault.

Moreover, since the reason of using wavelet packet transform is that it can extract information from raw signal at different frequency ranges, and from results, features 1 and 2 in Figure 7 (two highest normalized energies) are from low frequency domains. This further confirm that features from low frequency range can be selected in PEMFC diagnostics using polarization curve.

IV. CONCLUSION

The performance of polarization curves for on-line PEMFC diagnostics is investigated in the study. Since polarization curve can be measured efficiently at PEMFC system and can express various PEMFC losses, and by analyzing gradient evolution from measured curves, PEMFC performance could be monitored effectively.

With information extracted from measured curves, PEMFC condition could be easily monitored. Two approaches are applied in this paper, including the use of gradients from polarization curves, and utilization of normalized energy from wavelet packet transform. Based on findings, PEMFC system faulty state could be better discriminated with wavelet packet transform analysis. Moreover, the time-effectiveness of the analysis can meet the requirement of on-line fault diagnostic tasks.

Further study will be performed for investigating effectiveness of proposed method for PEMFC system using polarization curves containing more failure modes. Moreover, polarization curves including mass transport loss phase will be considered and included in the future diagnostic analysis.

REFERENCES

[1] R. Petrone, Z. Zheng, D. Hissel, M.C. Pera, C. Pianese, M. Sorrentino, M. Becherif, N. Yousfi-Steiner. "A review on model-based diagnosis methodologies for PEMFCs". Int. J. Hydrogen Energy , 2013, vol. 38, no. 17, pp. 7077-7091, 2013.

[2] Z. Zheng, R. Petrone, M.C. Pera, D. Hissel, M. Becherif, N.Y. Steiner, M. Sorrentino. "A review on non-model based diagnosis methodologies for PEM fuel cell stacks and systems". Int. J. Hydrogen Energy, vol. 38, no. 21, pp. 8914-8926, 2013.

[3] A. Ingimundarson, A.G. Stefanopoulou, D.A. McKay. "Model-based detection of hydrogen leaks in a fuel cell stack". IEEE Transactions on Control Systems Technology, vol. 16, no. 5, pp. 1004-1012, 2008.

[4] A. Forrai, H. Funato, Y. Yanagita, Y. Kato. "Fuel-cell parameter estimation and diagnostics". IEEE Transactions on Energy Conversion, vol. 20, no. 3, pp. 668-675, 2005.

[5] M.A. Rubio, A. Urquia, S. Dormido. "Diagnosis of performance degrdation phenomena in PEM fuel cells". Int. J. Hydrogen Energy, vol. 35, no. 7, pp. 2586-2590, 2010.

[6] L.A.M. Riascos, M.G. Simoes, P.E. Miyagi. "A bayesian network fault diagnostic system for proton exchange membrane fuel cells". Journal of Power Sources, vol. 165, no. 1, pp. 267-278, 2007.

[7] L. Placca, R. Kouta, D. Candusso, J.F. Blachot, W. Charon. "Analysis of PEM fuel cell experimental data using principal component analysis and multi linear regression". Int. J. Hydrogen Energy, vol. 35, no. 10, pp. 4582-4591, 2010.

[8] J. Hua, J. Li, M. Ouyang, L. Lu, L. Xu. "Proton exchange membrane fuel cell system diagnosis based on the multivariate statistical method". Int. J. Hydrogen Energy, vol. 36, no. 16, pp. 9896-9905, 2011.

[9] N.Y. Steiner, D. Hissel, Ph. Mocoteguy, D. Candusso. "Diagnosis of polymer electrolyte fuel cells failure modes (flooding & drying out) by neural networks modeling". Int. J. Hydrogen Energy, vol. 36, no. 4, pp. 3067-3075, 2011.

[10] S. Giurgea, R. Tirnovan, D. Hissel, R. Outbib. "An analysis of fluidic voltage statistical correlation for a diagnosis of PEM fuel cell flooding". Int. J. Hydrogen Energy, vol. 38, no. 11, pp. 4689-4696, 2013.

[11] N. Fouquet, C. Doulet, C. Nouillant, G.D. Tanguy, B.O. Bouamama. "Model based PEM fuel cell state-of-health monitoring via ac impedance measurements". Journal of Power Sources, vol. 159, pp. 905-913, 2006.

[12] Z. Zheng, M.C. Pera, D. Hissel, M. Becherif, K.S. Agbli, Y. Li. "A double-fuzzy diagnostic methodology dedicated to online fault diagnosis of proton exchange membrane fuel cell stacks". Journal of Power Sources, vol. 271, pp. 570-581, 2014.

[13] M.G. Santarelli, M.F. Torchio, P. Cochis. "Parameters estimation of a PEM fuel cell polarization curve and analysis of their behavior with temperature". Journal of Power Sources, vol. 159, no. 2, pp. 824-835, 2006.

[14] T.P. Ralph, M.P. Hogarth. "Catalysis for low temperature fuel cells". Platinum Metals Review, vol. 46, no. 1, pp. 3-14, 2002.

Suppression of Low-Frequency Magnetic Interference on Eddy Current Displacement Sensor with Nanometer Resolution

Chengliang Pan*
School of Instrument Science and Opto-electronics Engineering
Hefei University of Technology
Hefei, China
clpan@hfut.edu.cn

Shupeng Ren
School of Instrument Science and Opto-electronics Engineering
Hefei University of Technology
Hefei, China
1260466682@qq.com

Fei Yang
School of Instrument Science and Opto-electronics Engineering
Hefei University of Technology
Hefei, China
1669557838@qq.com

Chao Shi
School of Instrument Science and Opto-electronics Engineering
Hefei University of Technology
Hefei, China
695860065@qq.com

Heng Zuo
National Astronomical Observatories / Nanjing Institute of Astronomical Optics & Technology
Chinese Academy of Sciences
Nanjing, China
hengz@niaot.ac.cn

ZhiHua Feng
Department of Precision Machinery and Precision Instrumentation
University of Science and Technology of China
Hefei, China
fff@ustc.edu.cn

Abstract—Since eddy current displacement sensors (ECDSs) are based on phenomenon of electromagnetic induction, their measurement results would be seriously disturbed by the magnetic interference. Traditional solution is to add a shield integrated in the sensing probe, but it will also affect the property of sensing coil. In this paper, an alternative solution is proposed to suppress the low-frequency magnetic interference by high-frequency signal conditioning circuits with deliberate filtering segments. Prototype ECDS provides a resolution of 20 nm with measuring range of 50 µm and bandwidth of 3 kHz, under magnetic interference of 6000 A/m at 1 kHz.

Keywords—eddy current displacement sensor; magnetic interference; filter circuit; nanometer resolution

I. INTRODUCTION

High performance displacement sensors are extensively employed in the advanced scientific and industrial fields with precision positioning requirements [1-3], such as precision manufacturing, microscopic manipulation, and astronomical observation. These sensors are expected to be non-contact, high resolution, and immune to environmental interference [4]. The mostly utilized types are optical, capacitive, and eddy-current sensors [5, 6]. Among them, eddy current displacement sensors (ECDSs) are inherently compact, low cost, and tolerance of dirty environment, especially suitable for narrow displacement measurements in harsh working environments [7].

With the development of advanced sensing technologies, nanometer resolution has been achieved for the ECDSs [8, 9]. Meanwhile, as ECDSs are based on the phenomenon of electromagnetic induction, external magnetic fields would couple with the sensing coils [10], while the sensors are installed with power cables and electrical motors. Such interference will result in serious and intolerable error and noise of the sensors. Although the magnetic interferences on inductive sensors have been modeled to estimate their influence on the measurements [11], the general solution to eliminate the interference is to add a shielding layer [12]. However, the magnetic interferences in different frequency ranges require different shielding materials and the shielding

effects on sensing coils would be fatal for the precision displacement measurements [13]. Thus, other approaches should be developed to suppress the magnetic interference of ECDSs and remain the extremely high resolution .

In this paper, an ECDS with nanometer resolution is designed for the applications with magnetic interference. First, the influences of external magnetic field on the sensing coil of ECDS are estimated and the effects of the shielding layer are analyzed. Then, high-frequency signal conditioning circuits with deliberate filtering segments are proposed for amplification of weak displacement signal and suppression of magnetic interference. Finally, a prototype ECDS is manufactured and its performances under strong magnetic field interference are tested to confirm the proposed sensor.

II. INTERFERENCE AND SHIELD OF SENSING COIL

In general, a typical ECDS consists of a sensing coil, a conductive target, and corresponding signal conditioning circuit [14]. As shown in Fig. 1, an alternating current excites the sensing coil and generates an alternating magnetic field. The conductive target will induce a significant eddy current and generate an opposite magnetic field to the sensing coil. Meanwhile, the eddy current also dissipate electric energy due to the resistivity of conductive target. Such an interaction varies the equivalent inductance and resistance of the sensing coil. As the interaction statement is related to the distance between sensing coil and conductive target, precision relative displacement measurement can be obtained by demodulation of impedance variation with signal conditioning circuits.
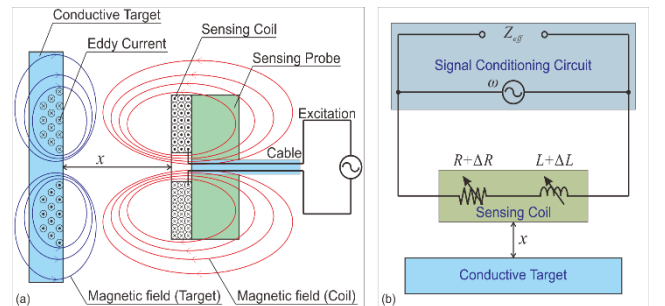
Fig. 1. Working principle (a) and equivalent circuit (b) of ECDSs.

*Chengliang Pan is the corresponding author. (E-mail: clpan@hfut.edu.cn).

Here, a finite element analysis model of ECDS is established to quantitatively investigate the influences of external magnetic field and shield. Figure 2 shows the 2D axisymmetric model established in FEA software COMSOL. It includes of a sensing coil, a conductive target, a shielding layer, and the surrounding air domain. The sensing coil has a 2 mm inner diameter, 5 mm outer diameter, and 0.3 mm thickness, with 120 turns copper wire of 50 μm diameter. The conductive target is an aluminum disk with 0.5 mm thickness and 30 mm diameter. The shielding layer is a SUS304 tube with 6 mm inner diameter, 7 mm outer diameter, and 5 mm length. The material parameters of the components are listed in Table 1.
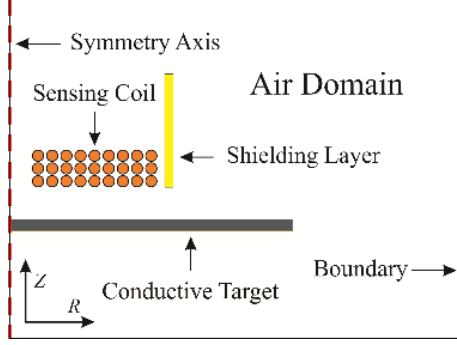
Fig. 2. FEA model of ECDS.

TABLE I. MATERIAL PARAMETERS OF THE ECDS MODEL

Material	Conductivity (MS/m)	Relative Permeability
Copper	60.0	1
Aluminum	37.7	1
SUS304	1.03	1

According to the Faraday's law, the induced voltage on the sensing coil is proportional to the intensity and frequency of the external alternating magnetic field. Without shielding layer, 1 kHz magnetic fields along the Z-axis and R-axis are respectively applied in the air domain with intensity from 100 A/m to 1000 A/m. Figure 3 shows the coupled voltage E on the sensing coil along with the intensity H of external magnetic field. The coupled voltage E is sensitive to the magnetic field along the Z-axis. The amplitude of coupled voltage reaches up to 100 mV under the external alternating magnetic field of 8000 A/m at 1 kHz. The coupled voltage is insignificant to the magnetic field along the R-axis as the common sense.
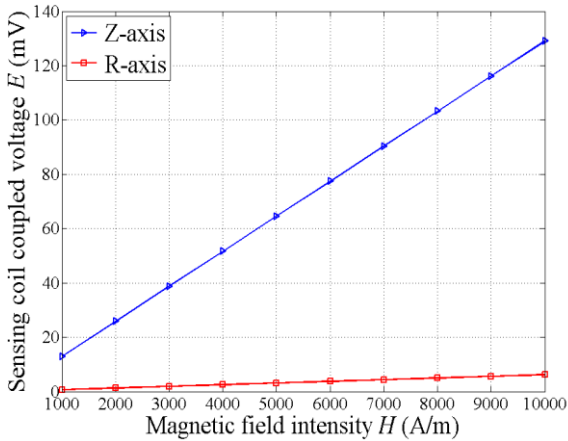
Fig. 3. Coupled voltage of sensing coil with magnetic field at 1 kHz.

When the shield is added in the model, magnetic flux will redistribute with an obvious variation. Figure 4 demonstrates the magnetic flux density of the ECDS without and with shielding layer, while a 1 V, 1 MHz voltage is applied on the sensing coil. It can be seen that the magnetic flux generated by the sensing coil is blocked and cannot penetrate the shielding layer. Since the magnetic flux is restrained in the shielding tube, it directly affects the eddy current distribution in the conductive target, and provides a great impact on the equivalent impedance of the sensing coil.
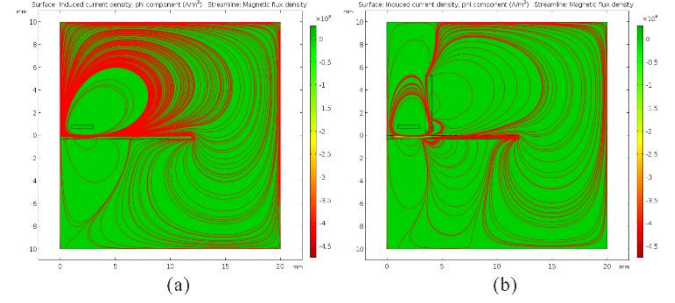
Fig. 4. Magnetic fields of the ECDS (a) without shield (b) with shield.

Figure 5 shows the inductance and resistance of the sensing coil at 1 MHz without and with the shielding layer. After adding the shielding layer, it is clear that the inductance of sensing coil decreases, and the farther the distance is, the more the decrease is, which means the sensitivity of the sensing coil is greatly decreased. At the same time, variation trend of the sensing coil resistance arises a dramatic change, corresponding to a more complex nonlinear relationship. Thus, the sensitivity and accuracy of the ECDSs would be seriously affected by the adding shield. In addition, the shield will also increase the manufacturing difficulty and production cost of the sensing probe.
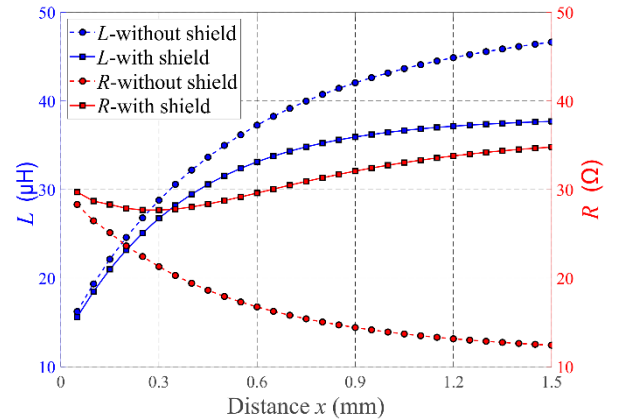
Fig. 5. Impedance variations of the sensing coil along with distance of conductive target.

III. SIGNAL CONDITIONING CIRCUIT

Based on the simulation results of the sensing coil, it is considered that shield design of the sensing coil is not a perfect solution to reject of the interference of external magnetic field. Since the magnetic interferences in the industrial plants and scientific facilities mainly come from the high current cables and electromagnetic motors, their significant fringe fields usually appear as slowly varying magnetic fields. These low-frequency interference signals coupled in the sensing coil would be suppressed to an acceptable faint value with deliberate filtering segments among the high-frequency signal conditioning circuit.

A. Basic Circuit

The AC bridge amplification and modulation circuit is a typical ESDC signal conditioning circuit with high resolution and stability [15]. As shown in Fig. 6, the basic circuits consist of a sine-wave generator, a well-matched AC bridge, a pre-amplifier, a lock-in amplifier, and an output amplifier.
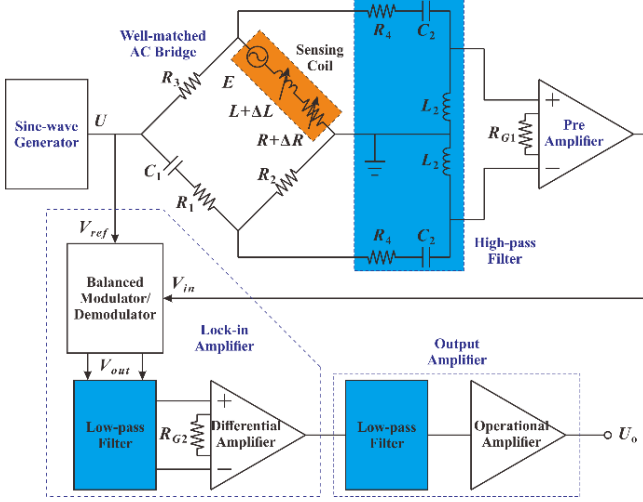
Fig. 6. Schematic of the signal conditioning circuit of the ECDS.

A novel AC bridge is adopted to extract the impedance variation of the sensing coil [16]. A capacitor C_1 and resistor R_1 series branch is used to replace the traditional reference coil, locating in the opposite arm of the sensing coil. Other two arms are resistor R_2 and R_3. Such a bridge could avoid the influences of high temperature drift on the resistance of the traditional reference coil. Replacing all the components with high precision and low temperature drift electronic elements, the stability of the bridge can be greatly improved.

While the angular frequency of the exciting voltage U is ω, the balanced condition of the bridge can be expressed as

$$\frac{j\omega L + R}{R_3} = \frac{j\omega C_1 R_2}{j\omega C_1 R_1 + 1} \tag{1}$$

Since the relative impedance variation of the sensing coil is set very small in the measuring range, as

$$\Delta R \ll R, \quad \Delta L \ll L \tag{2}$$

the differential mode voltage of the bridge output signal can be approximately calculated as

$$\Delta U = \frac{R_3(\Delta R + j\omega_1 \Delta L)}{(R_3 + R + j\omega L)^2} U \tag{3}$$

Here, R_3 appears a great influence on the sensitivity and linearity of the sensor. Increasing the value of R_3 would improve the linearity, but at the cost of sensitivity loss. Therefore, an optimal fixed value of R_3 is needed. Other components can be calculated from the following formulas

$$R_1 = \frac{R}{\omega^2 L C_1}, \quad R_2 = \frac{L + RR_1C_1}{R_3C_1} \tag{4}$$

The bridge is excited by a 1 MHz sinusoidal voltage from the sine-wave generator (AD9850 DDS module with THS4031 low-noise, high-speed power amplifier). The values of the sensing coil (includes co-axial cable) and other components of the balanced bridge are listed in Table 2.

TABLE II.   VALUES OF THE BALANCED AC BRIDGE COMPONENTS

L (µH)	R (Ω)	C_1 (pF)	R_1 (Ω)	R_2 (Ω)	R_3 (Ω)
35	25	725	25	105	470

A pre-amplifier (INA103, low noise, low distortion instrumentation amplifier) is used to amplify the differential mode voltage of the bridge output signal. Then a lock-in amplifier (MC1496 balanced modulator/demodulator with low-pass filter and INA103 amplifier) is used to demodulate the displacement signal. Finally, the displacement signal is amplified to the desired magnitude with the output amplifier. A total gain of displacement signal is about 1000 among the whole signal conditioning circuit.

B. Filtering Segments

For the proposed sensor, the band width of the sensing displacement is set to 3 kHz, and the band width of the magnetic interference is set to 1 kHz. From the simulation results, the interference signal E coupled in the sensing coil would be up to hundred millivolts. Without preconditioning of the disturbed bridge output, the output of the pre-amplifier would be saturated by E. Here, a RLC high pass Butterworth filters is added before the pre-amplifier. To avoid the adverse influence on the balance of AC bridge, the RLC filters are symmetrically set and the values of RLC components are set to 210 Ω, 220 µH, and 10 nF, with a corner frequency of 103 kHz. Bode diagram of the high pass filter is shown in Fig. 7. The interference signal at 1 kHz would be suppressed to -80 dB magnitude, which means the coupled voltage is less than 0.01% of the primary value. Meanwhile, the magnitude of sensing signal around 1 MHz does not have significant attenuation.
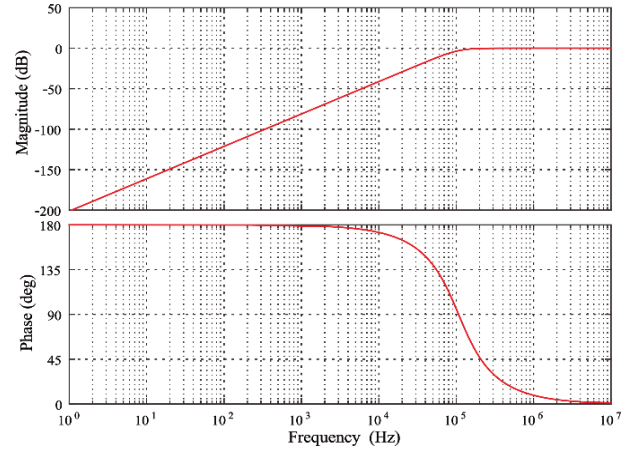
Fig. 7. Bode diagram of RLC high pass Butterworth filter.

After the balanced demodulator, the sensing signal is demodulated from the high-frequency band to low-frequency band. Low pass filter is needed to filter their residual higher harmonics. Meanwhile, the interference signal is also amplified by the pre-amplifier and modulated from the low-frequency band to high-frequency band. To restrain the magnitude of the interference signal, four-order low pass Butterworth filters are used in the lock-in amplifier, as shown in Fig. 8. For further suppression of the noise, such low pass filter is also applied in the output amplifier. Bode diagram of the low pass filters is shown in Fig. 9. The corner frequencies of the filters in the lock-in amplifier and output amplifier are set to 25 kHz and 3 kHz, respectively. The interference signal at 1 MHz would be suppressed to -190 dB and -130 dB magnitude, respectively. Meanwhile, the magnitude of

sensing signal below 3 kHz does not have significant attenuation.
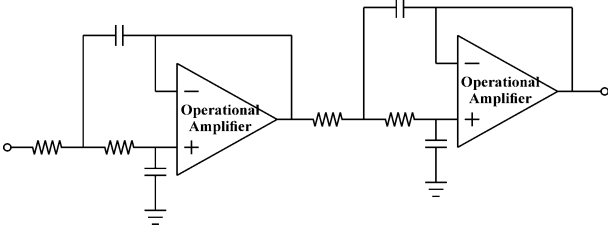
Fig. 8. Four-order low pass Butterworth filter.
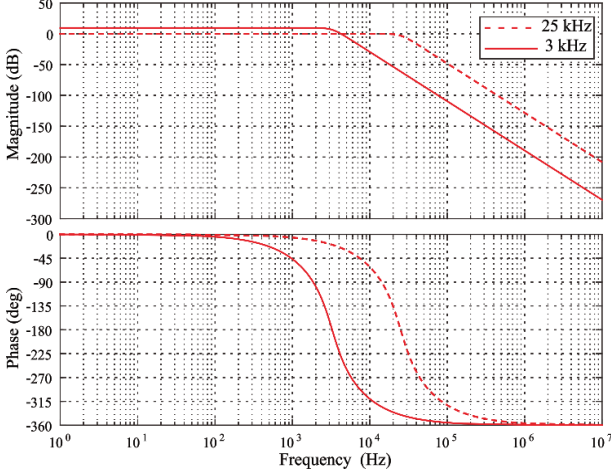
Fig. 9. Bode diagram of four-order low pass Butterworth filters.

With deliberate filtering segments among the AC bridge amplification and modulation circuit, the interference signal of several hundred millivolt would be suppressed to several millivolt in the output voltage, which is a comparable level of the circuit basic noise. The magnitude attenuation of sensing signal is negligible, but the phase shifting of the RLC filter at 1 MHz will affect the gain of the lock-in amplifier and the phase shifting of the low pass filters will influence the signal delay of the vibration measurement in the high-frequency band.

IV. EXPERIMENTS AND RESULTS

A prototype ECDS is manufactured for the experimental investigations of the magnetic interference and the solution of filter circuits. The sensing coil has the same parameter as shown in the FEA simulation. It is attached on a quartz rod as the sensing probe and connected to the signal conditioning circuit with a co-axial cable. With ± 10V output voltage in the 50 μm measurement range of 300 μm balanced distance, the gains of the pre-amplifier, lock-in amplifier, output amplifier are set to about 118, 2.5, and 3, respectively. A voice coil is used to generate the simulated magnetic field. And the performances of the prototype ECDS are tested under the strong magnetic field interference.

A. Voice Coil

The voice coil has 8 mm inner diameter, 20 mm outer diameter, and 14 mm height, with about 130 turns copper wire of 0.8 mm diameter. Figure 10 shows the magnetic field distribution of the voice coil at 1 kHz, which appears a concentrated magnetic flux cross the sensing coil. The intensity of the magnetic field is about 6000 A/m with 0.5 A exciting current on the voice coil. The coupled voltage E on the sensing coil reach up to 90 mV, which is a great quantity compared with the normal differential voltage of the AC bridge output signal.
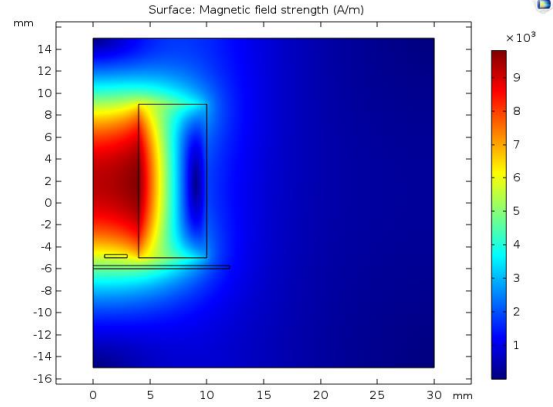
Fig. 10. Magnetic field of voice coil.

B. ECDS Linearity and Resolution

The linearity of the prototype ECDS is tested with a precision nanometer stage (PI P-620.1CD with controller E625.CR) and a digital multimeter (Fluke 8845A). The stage provides step displacement of 1 μm in the measurement range, and the output voltages of the prototype ECDS are recorded by the multimeter. Figure 11 shows the relationship between the output voltage of the prototype ECDS and the displacement of the nanometer stage. In the whole measurement range of 50 μm, the non-linearity of the prototype ECDS is better than 0.5% without any compensation. The basic noise of the prototype ECDS is about 10 mV$_{p-p}$ in the bandwidth of 3 kHz, corresponding to the displacement resolution of 4 nm. As shown in Fig. 12, the prototype ECDS can clearly distinguish the 5 nm step displacements at 1 Hz generated by the PI nanometer stage.
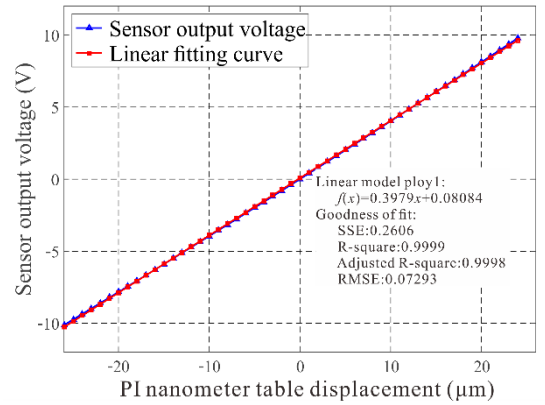
Fig. 11. Linearity of the prototype ECDS.
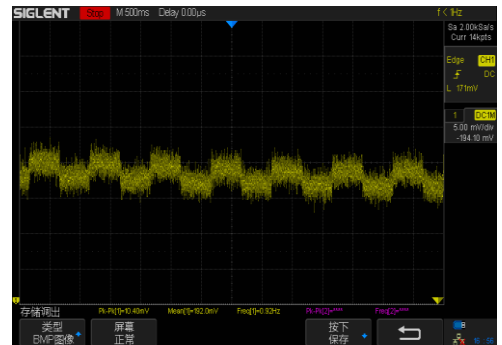
Fig. 12. Output voltage with 5 nm step displacements.

C. Strong Magnetic Field Interference

Figure 13 shows the experiment platform used to test the performances of the prototype ECDS under the magnetic interference of the voice coil. The voice coil is fixed on the base and driven by an OPA541 power amplifier. The sensing probe is coaxially set cross the inner hole of the voice coil. The conductive target is attached on the central lower surface of a steel beam, while the two ends of the steel beam are fixed. A piece of PZT sheet is attached to the central upper surface of the steel beam, as the actuator of the beam vibration driven by a PA 94 power amplifier. The output of the prototype ECDS is compared with the result of a laser displacement sensor (Micro-Epsilon optoNCDT 2300).
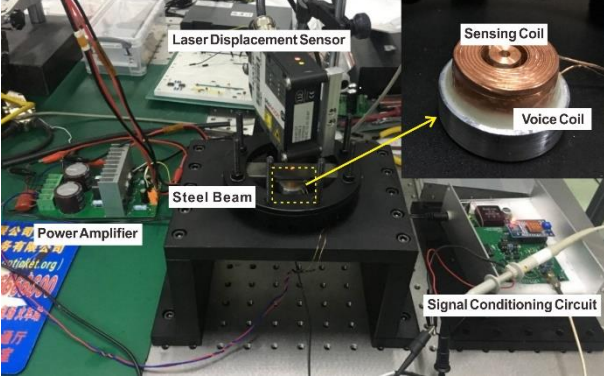
Fig. 13. Test platform of strong magnetic field interference.

Figure 14 shows the voltages of the signal conditioning circuits under the magnetic interference. The voice coil is applied with a 0.5 A exciting current at 1 kHz. The sensing coil is coupled with a 100 mV_{p-p} interference signal at the output point of the AC bridge (Fig. 14a). With the suppression of RLC high pass Butterworth filters, the output voltage of the pre-amplifier becomes clear (Fig. 15b), while the interference signal is greatly suppressed, even amplified 118 times by the pre-amplifier. With the lock-in amplifier, the high-frequency sensing signal is demodulated to low-frequency band. The signal is dealt with the fourth-order low pass Butterworth filter with a corner frequency of 25 kHz (Fig. 14c). With the following differential amplifier and output amplifier of 7.5 times gain, the interference signal are further eliminated into about 60 mV_{p-p} in the output voltage (Fig. 14d), which means a resolution of 20 nm is remained.
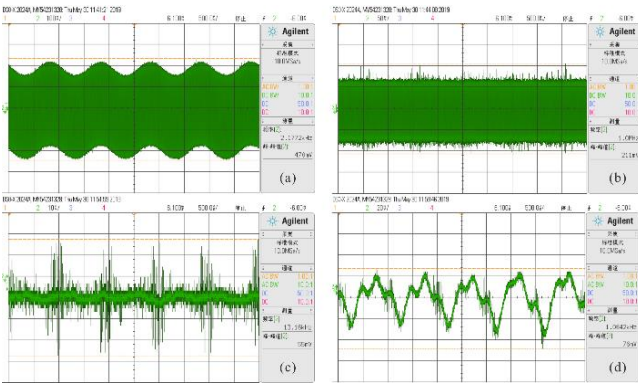
Fig. 14. Voltages of the signal conditioning circuit under the magnetic interference.

Figure 15 shows the vibration amplitude of the steel beam excited by the PZT sheet at its resonant frequency of 730 Hz and suffered magnetic interference of the voice coil. Compared with the laser displacement sensor, the measured values of prototype ECDS remain a good linearity in the full-scale range of 50 μm.
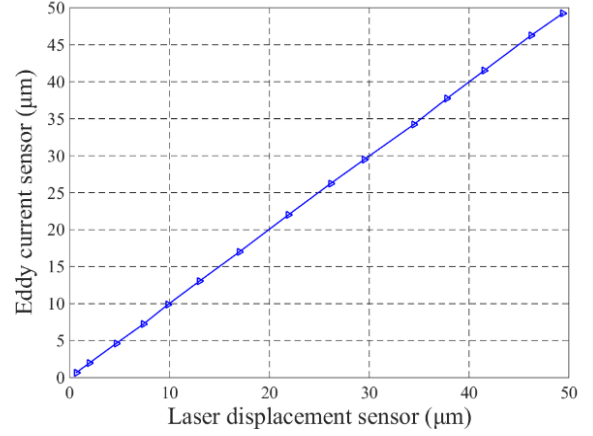
Fig. 15. Comparsion of prototype ECDS and Laser displacment sensor.

V. CONCLUSIONS

In this paper, an ECDS used for strong magnetic field environment is designed, replacing the traditional solution of shielding method. The influences of external strong magnetic field are estimated to be critical to the measurement accuracy of ECDSs, and the effectiveness of shielding layer on the sensing coil are limited for the precision measurement. New signal conditioning circuits based on the AC bridge amplification and modulation circuits and deliberate filtering segments are designed to amplify the displacement signal and suppress the magnetic interference simultaneously. Experiment results indicate the nanometer resolution and good linearity of the prototype ECDS, even under the strong magnetic field environment. With less increases of the design complexity and manufacturing cost, the proposed ECDS provides a great practical significance and comprehensive measurement ability for complex applications.

REFERENCES

[1] W. L. Pearn and C. S. Chang, "Precision measures for processes with multiple manufacturing lines," Int. J. Adv. Manuf. Technol., vol. 30, pp. 1202-1210, 2006.

[2] C. Y. Shi, D. K. Luu, Q. M. Yang, J. Liu, J. Chen, C. H. Ru, S. R. Xie, J. Luo, J. Ge, and Y. Sun, "Recent advances in nanorobotic manipulation inside scanning electron microscopes," Microsyst. Nanoeng., vol. 2, pp. 16024, 2016.

[3] H. Zuo, G. P. Li, and C. L. Pan, "Non-contact displacement measure method based on eddy current sensors in the large aperture adaptive mirror system," Proceedings of SPIE, vol. 10703, pp. 107037B, 2018.

[4] A.J. Fleming, "A review of nanometer resolution position sensors: Operation and performance," Sens. Actuators A Phys., vol. 190, pp. 106-126, 2013.

[5] P. J. de Groot, "A review of selected topics in interferometric optical metrology," Rep. Prog. Phys., vol. 82, pp. 056101, 2019.

[6] B. George, Z. C. Tan, and S. Nihtianov, "Advances in capacitive, eddy Current, and magnetic displacement sensors and corresponding interfaces," IEEE Trans. Ind. Electron., vol. 64, pp. 9595-9607, 2017.

[7] Eddy-current displacement sensors overview, Application Notes, Lion Precision, 2011.

[8] H. B. Wang and Z. H. Feng, "Ultrastable and highly sensitive eddy current displacement sensor using self-temperature compensation," Sens. Actuator A Phys., vol. 203, pp. 362-368, 2013.

[9] Z. L. Qu, Q. Zhao, and Y.G. Meng, "Improvement of sensitivity of eddy current sensors for nano-scale thickness measurement of Cu films," NDT E. Int., vol. 61, pp. 53-57, 2014.

[10] L. Zhang, C. Deng, X. M. Zhang, J. X. Li, and J. Meng, "Research of electromagnetic interference on the eddy current sensor in the inverter-motor driving system," Proceedings of 2016 Asia-Pacific International Symposium on Electromagnetic Compatibility (ISEC 2016), 1175-1177, Shenzhen, P.R. China, 2016.

[11] M. Martino, G. Golluccio, R. Losito, and A. Masi, "An analytical model of the effect of external DC magnetic fields on the AC voltages of an LVDT," Proceedings of IEEE International Instrumentation & Measurement Technology Conference (I2MTC 2010), pp. 213-218, Austin, USA, 2010.

[12] A. Masi, M. Lamberti, R. Losito, and M. Martino, "Reduction of magnetic interference on the position sensors of the LHC collimators," Proceedings of 2nd International Particle Accelerator Conference (IPAC 2011), pp. 1623-1625, San Sebastián, Spain, 2011.

[13] Z. H. Liu, J. F. Yao, C. F. He, L. M. Li, X. C. Liu, and B. Wu, "Development of a bidirectional-excitation eddy-current sensor with magnetic shielding: detection of subsurface defects in stainless steel," IEEE Sens. J., vol. 18, pp. 6203-6216, 2018.

[14] H. B. Wang, W. Li, and Z. H. Feng, "A compact and high-performance eddy-current sensor based on meander-spiral coil," IEEE Trans. Magn., vol. 51, pp. 4003806, 2015.

[15] M. R. Nabavi and S Nihtianov, "Eddy-current sensor interface for advanced industrial applications," IEEE Trans. Ind. Electron., vol. 58, pp. 4414-4423, 2011.

[16] H. B. Wang, B. Ju, W. Li, and Z. H. Feng, "Ultrastable eddy current displacement sensor working in harsh temperature environments with comprehensive self-temperature compensation", Sens. Actuator A Phys., vol. 221, pp. 98-104, 2014.

Vibration Feature of Locomotive Axle Box with a Localized Defect in its Bearing Outer Race

Yuqing Liu
State Key Laboratory of Traction Power
Southwest Jiaotong University
Chengdu, China
380562002@qq.com

Zaigang Chen*
State Key Laboratory of Traction Power
Southwest Jiaotong University
Chengdu, China
zgchen@home.swjtu.edu.cn

Abstract—**The axle box bearing is a critical component to support the vehicle and reduce the friction effect between the wheelset and the axle box. As one of the major failure modes, the localized defect in the outer race of the axle box bearing is a great threat to the operation safety of railway vehicles. Therefore, it is helpful to investigate the dynamic features of locomotive with effect of a localized defect of axle box bearing so as for better understanding of the bearing fault mechanism. In this paper, a locomotive-track coupled dynamics model with traction transmissions is proposed where the dynamics sub-model of the axle box rolling bearing is included. And the impact effect between the roller and the defect region is described as the time-varying displacement excitation and the time-varying contact stiffness excitation of the fault bearing. The results show that the passing frequency of roller and its harmonics of the vertical acceleration vibration of the axle box could be good indicators for the diagnosis of this localized defect, even if the fault feature could not be observed in the time history signals under the wheel-rail interaction. This work can provide some guidance for the early localized defect detection and diagnosis for the axle box bearing.**

Keywords—localized defect, axle box bearing, gear transmission, locomotive-track coupled dynamics, failure detection

I. INTRODUCTION

Axle box bearing is a key component of the locomotive to support the vehicle and reduce the friction interaction between the wheel axle and the axle box. The localized defect will propagate around the corresponding race along the roller motion direction until catastrophic failure occurs. Therefore, the localized defect is potentially dangerous for the operation stability and safety of the locomotive. It is helpful to investigate the dynamic responses of the locomotive, especially the axle box, with effect of a localized defect of axle box bearing for the bearing fault detection and diagnosis.

A large number of researches have been focused on the vibration responses of the rolling bearing under the excitation from the localized defect. For instance, Liu et al. [1,2] performed a series of investigations on the fault mechanism of effect of the localized defect of the bearing. Considering the time-varying displacement excitation, time-varying contact stiffness excitation and additional deformations at the sharp edges of the defect, they established a lumped spring-mass dynamics model of rolling bearing to investigate the fault mechanism of the localized defect. In addition, Liu [3] further proposed a dynamic modeling method of a rotor-roller bearing-housing system, which can provide accurate simulation results to detect the localized defect in this system. Li et al. [4] paid attentions to the effect of the localized defect

on the contact ball bearing. Based on the finite element method and the lumped mass method, Yang et al. [5] analyzed the dynamic responses of the rotor-rolling bearing-casing system through the dynamic simulation and experimental analysis.

However, the dynamic characteristics of the axle box bearing with a localized defect on the outer race has not been reflected from the previous researches, which is attributed to the neglect of interactions between the rolling bearing system and the locomotive-track coupled dynamics system. Fortunately, development of the railway vehicle dynamics models has supplied good possibilities. For example, Chen et al. [6,7] proposed a locomotive-track vertical\vertical-longitudinal coupled dynamics model with gear transmissions to obtain the accurate internal excitation from the transmission systems, such as the time-varying mesh stiffness and dynamic transmission error. Wang et al. [8,9] performed some investigations on the dynamic responses of the axle box bearing with the effect of high-speed train vibration environment. And Li et al. [10] established a vehicle-track spatially coupled dynamics model considering the nonlinear contact forces of the axle box bearing. Based on the multi-feature parameters, Li et al. [11] introduced a fault diagnosis method for the axle box bearing.

The previous researches provide a sound theoretical basis for the dynamic investigation of locomotive axle box bearing with a localized defect. The main objective of this paper is to use established the locomotive-track coupled dynamics model with traction transmissions for revealing the dynamic features of a localized defect in the outer race of the axle box bearing. And the internal excitations excited from the gear transmission subsystem and axle box\motor bearing subsystems, the nonlinear wheel-rail interaction and the roller-defect region impact effect are fully considered in this model. The simulation results can provide some guidance for the early defect detection and diagnosis for the axle box bearing.

II. LOCOMOTIVE-TRACK COUPLED DYNAMICS MODEL WITH A LOCALIZED DEFECT IN AXLE BOX BEARING

To analyze the dynamic responses of the locomotive under the internal impact force excited by the localized defect of the axle box bearing, based on the locomotive-track coupled dynamics model with gear transmissions which was proposed by Chen et al. in Ref. [7], a locomotive-track coupled dynamics model with traction transmissions is proposed and the internal interactions between the components of the axle box bearing are considered in detail. As shown in Fig.1, the locomotive-coupled dynamics model consists of the vehicle subsystem, the track structure subsystem, the gear transmission subsystem and the rolling bearing subsystem.

In the vehicle subsystem, one car body, two bogie frames and four wheelsets are contacted via the primary and the secondary suspensions, respectively. The traction torque is transmitted

from the traction motor to the wheelset through the gear transmission and the longitudinal motion of the locomotive is driven by the tractive forces generated from the wheel-rail interface. An improved mesh stiffness calculation model of spur gear pair with tooth profile deviations [12] is applied to calculate the accurate time-varying mesh stiffness of the gear pair of the traction transmission. In addition, the classical ballasted track structure is used in this dynamics model, which is composed of the rail, rail pads, sleepers, ballasts and subgrade.
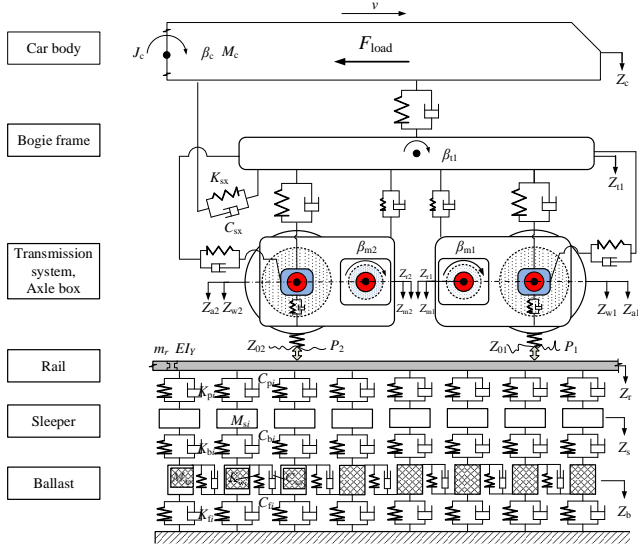
Fig. 1. Locomotive-track coupled dynamics model with gear transmissions.

In order to reflect the impact effect between the roller and the defect region, a dynamics model of rolling bearing considering contact forces between each component and the corresponding friction forces is proposed in this paper. Here, the inner ring is fixed on the wheelset and the outer ring rigidly connects with the axle box. The time-varying displacement excitation and the time-varying stiffness excitation of the failure bearing can be described by an improved analytical model proposed by Liu et al. in Ref. [1].
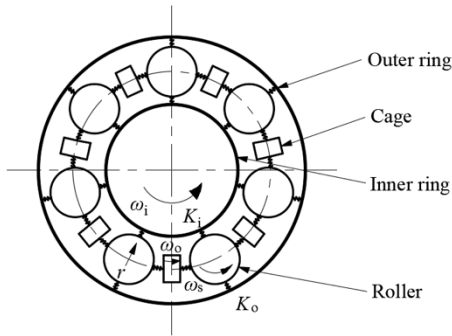
Fig. 2 A dynamics model of the axle box bearing.

When a localized defect of the bearing is in the middle or late stage of plastic deformations at the propagated defect edges, the schematic of the contact relationship between the roller and the defect region are shown in Fig.3. It can be seen that there are some visible differences of the relative radial displacement and the contact stiffness between the roller and the outer race. Therefore, the impact interaction between the roller and the defect region can be described as the comprehensive effect of the time-varying displacement

excitation and the time-varying contact stiffness excitation induced by the localized defect on the outer race.
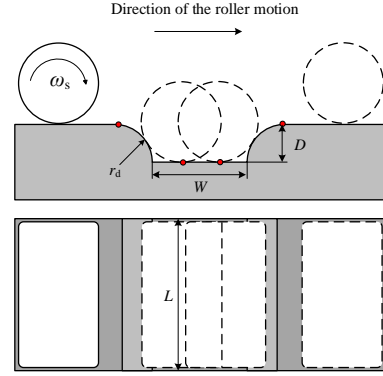
Fig. 3 Schematic of the contact relationship between the roller and the middle\late localized defect.

For a localized defect with the cylindrical surface edges, the time-varying displacement excitation of the localized defect can be represented as a piecewise function [13]:

$$
H = \begin{cases}
D\sin\left(0.5\pi / \Delta T_1 \left(\text{mod}\left(\theta_{oi}, 2\pi\right)\right)\right) & \theta_0 \leq \text{mod}\left(\theta_{oi}, 2\pi\right) \leq \theta_1 \\
D & \theta_1 \leq \text{mod}\left(\theta_{oi}, 2\pi\right) \leq \theta_2 \\
D\sin\left(0.5\pi / \Delta T_2 \left(\text{mod}\left(\theta_{oi}, 2\pi\right)\right)\right) & \theta_2 \leq \text{mod}\left(\theta_{oi}, 2\pi\right) \leq \theta_3 \\
0 & \text{otherwise}
\end{cases}
\tag{1}
$$

where D is the depth of the localized defect. θ_{oi} represents the rotational angular displacements of ith roller. $\theta_0\backslash\theta_3$ and $\theta_1\backslash\theta_2$ are the angular positions of the start and end points of the localized defect and the roller-bottom surface contact area, respectively. ΔT_1 and ΔT_2 can be calculated by:

$$
\Delta T_1 = \Delta T_2 = \left(r_d + \sqrt{\left(R_r + r_d\right)^2 - R_r^2}\right) / R_o
\tag{2}
$$

where R_r and R_o are the radiuses of the roller and the outer race, respectively. r_d is the radiuses of the cylindrical surface edge.

According to the Hertz contact theory, the contact stiffness between the roller and the cylindrical surface at the localized defect edge can be deduced by [14]:

$$
K_{rd} = \frac{dQ}{d\delta_{rd}}
\tag{3}
$$

where Q is the external force. δ_{rd} is the relative elastic deformation between the roller and the edge, which can be calculated by:

$$
\delta_{rd} = \frac{2Q}{\pi l_e} \frac{1-\upsilon^2}{E_0} \left(\ln \frac{4R_r r_d}{b_{id}^2} + 0.814 \right)
\tag{4}
$$

where l_e is the equivalent length, υ and E_0 denote the Poission's ratio and equivalent elastic modulus. And bid is the semi-width of the contact surface.

The contact stiffness between the roller and the bottom surface of the localized defect can be calculated by [15]:

$$
K_b = \frac{dQ}{d\delta_b}
\tag{5}
$$

where δ_b is the relative elastic deformation between the roller and the bottom surface, which can be calculated by:

$$
\delta_b = \frac{4R_r^{0.5}}{3} \left(\frac{1-\upsilon_r^2}{E_r} + \frac{1-\upsilon_o^2}{E_o} \right)
\tag{6}
$$

where $\upsilon_r\backslash\upsilon_o$ and $E_r\backslash E_o$ denote the Poission's ratio and equivalent elastic modulus of the roller and the outer ring, respectively.

Considering the large degrees of freedom of the dynamics model, an explicit-implicit hybrid method is proposed here to achieve the desired calculation speed and accuracy. The fast-explicit integration method [16] is utilized to solve the dynamic equations of the vehicle, the track structure and the gear transmission subsystems. While for the bearing subsystem, a fourth-order Runge–Kutta method [17] is utilized to meet the precision requirements. The integration steps of two numerical methods are 3×10^{-5} s and 5×10^{-6} s, respectively.

III. DYNAMIC INVESTIGATION AND RESULT DISCUSSIONS

To reveal the vibration features of the locomotive axle box in presence of local defect in its outer race, a HX locomotive is used for dynamic simulation in this paper, and its major design parameters can be referenced to Ref. [7]. The locomotive running speed is set as 80 km/h and the traction torque of the motor is 1.5 kN/m. To consider the changes in the contact zones near the defect edges, the defect length is assumed to be larger than the roller length and its width is less than the roller diameter. Thus, the defect length, width and depth are assumed to be 14 mm, 5 mm and 0.1 mm, respectively. And the angular displacement of the start of the localized defect (θ_0) is $\pi/2$. In addition, the characteristic frequencies of the axle box are listed in Table 1.

TABLE I.   CHARACTERISTIC FREQUENCIES OF THE AXLE BOX

Cage frequency f_c (Hz)	Roller passing outer race frequency f_b (Hz)	Mesh frequency f_m (Hz)
2.551	73.988	679.061
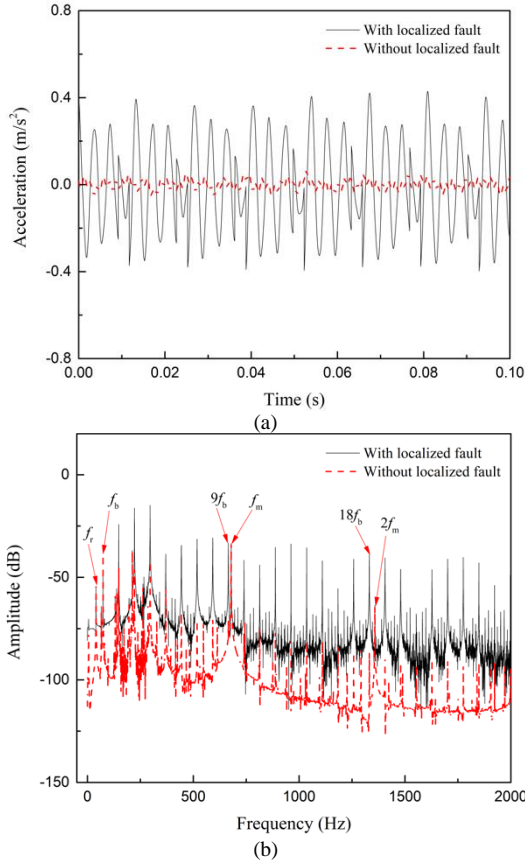
Fig. 4 Vertical accelerations of the axle box without track irregularity: a) time histories and b) frequency spectrum.
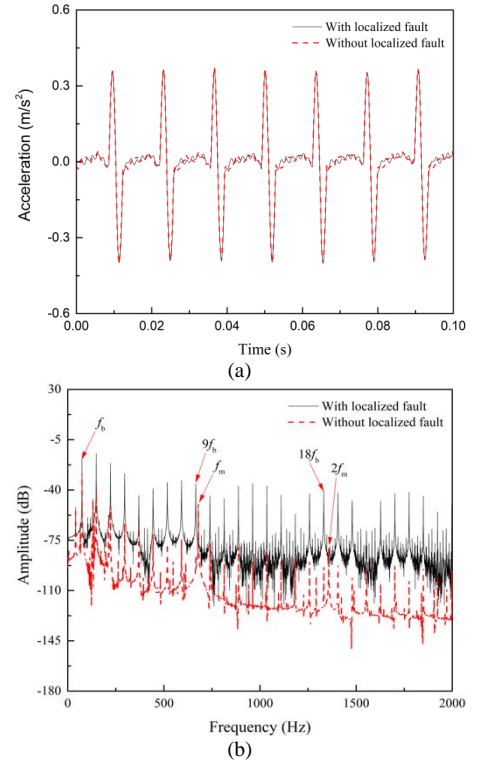
Fig. 5 Longitudinal vibration accelerations of the axle box without track irregularity: a) time histories and b) frequency spectrum

To highlight the effect of the localized defect of the axle box, time histories and frequency spectrum of the vertical and longitudinal accelerations of the axle box are extracted from the simulation without considering the track irregularity. As shown in Figs. 4 and 5, it can be seen that the peak-peak value of the vertical acceleration increases evidently under the excitation of the defect, while the localized defect has an inconspicuous effect on the axle box in longitudinal direction in time histories. After analyzing the spectrum of the accelerations, the frequency of roller passing the outer race and its harmonics are more evident than the mesh frequency of the gear transmission while there is no obvious change in longitudinal direction. In addition, it should be noted that there are frequency bands of 16 Hz and 29 Hz around the passing frequency of roller (PFOR) and its harmonics. The similar phenomenon also can be observed in Fig. 5(b). Under the effect of the gravity of the locomotive and the interaction between the wheelset and the rail, the load region of the axle box bearing is formed on the top of the outer race. Therefore, the impact force between the roller and the defect region mainly affects the dynamic responses of the axle box in vertical direction while the changes of longitudinal acceleration are weak.
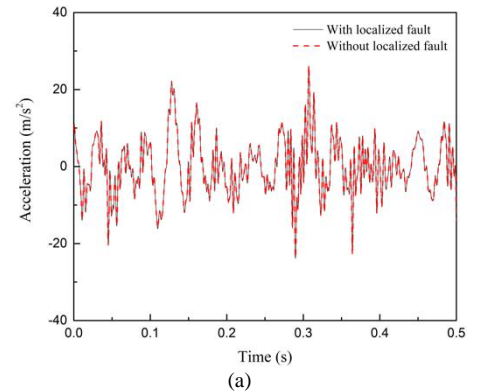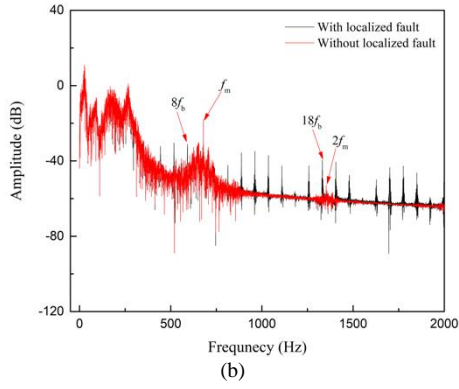
Fig. 6 Vertical accelerations of the axle box with track irregularity: a) time histories and b) frequency spectrum
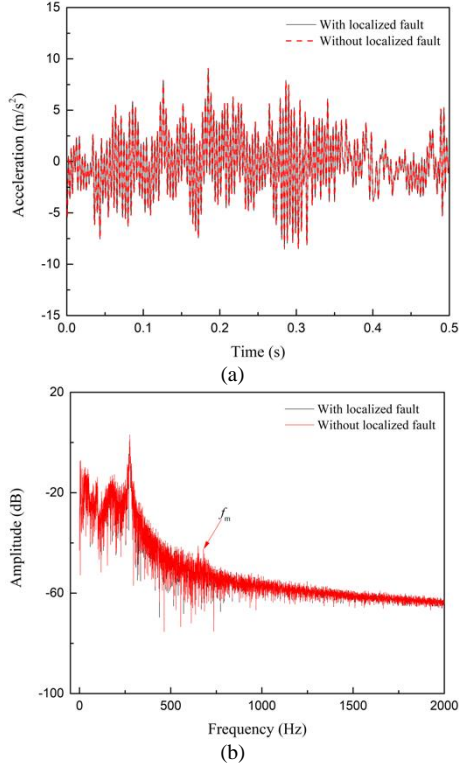
Fig. 7 Longitudinal vibration accelerations of the axle box with track irregularity: a) time histories and b) frequency spectrum

Figures 6 and 7 illustrate the effect of the localized defect of the axle box under the external wheel-rail interaction in actual operation of the locomotive. In order to truly reflect the railway conditions, the Association of American Railroads 5 class irregularity is adopted here. It can be seen that there are no significant changes in the time histories of the vertical and longitudinal accelerations. In addition, the frequency of roller passing the outer race and its harmonics become overwhelmingly obvious. The frequency bands of 16 Hz are covered and them of 29 Hz are not apparent under the external excitation from the wheel-rail interface. However, considering the effect of the track irregularity, there is no clear characteristic signal extracted from the dynamic represents of the axle box in the longitudinal direction, which means that it loses value for the localized defect diagnosis to analyze longitudinal acceleration.

IV. CONCLUSIONS

A locomotive-track coupled dynamics model with traction transmissions is established in this paper by considering the time-varying mesh stiffness and the internal excitation of the rolling bearing, where the impact effect of the localized outer race defect of the axle box bearing is represented by the time-varying displacement excitation and time-varying contact stiffness excitation. And then, dynamic simulations are performed to receive the vibration responses of the locomotive dynamic system, and the vibrations of the axle box are extracted for analysis under the complicated excitations. The results indicate that the localized outer race defect of the axle box bearing is mainly reflected in the vertical vibration accelerations. The passing frequency of rollers and its harmonics are the key characteristic signals, and the frequency bands around the passing frequency of rollers and its harmonics can be a good indicator for the axle box bearing fault detection.

REFERENCES

[1] Liu J, Shao Y. An improved analytical model for a lubricated roller bearing including a localized defect with different edge shapes. Journal of Vibration and Control, 2018, 24(17): 3894-3907.

[2] Liu J, Shao Y. Dynamic modeling for rigid rotor bearing systems with a localized defect considering additional deformations at the sharp edges. Journal of Sound and Vibration, 2017, 398: 84-102.

[3] Liu J. A dynamic modelling method of a rotor-roller bearing-housing system with a localized fault including the additional excitation zone. Journal of Sound and Vibration, 2020, 469: 115144.

[4] Li X, Yu K, Ma H, et al. Analysis of varying contact angles and load distributions in defective angular contact ball bearing. Engineering Failure Analysis, 2018, 91: 449-464.

[5] Yang Y, Yang W, Jiang D. Simulation and experimental analysis of rolling element bearing fault in rotor-bearing-casing system. Engineering Failure Analysis, 2018, 92: 205-221.

[6] Chen Z, Zhai W, Wang K. A locomotive–track coupled vertical dynamics model with gear transmissions. Vehicle System Dynamics, 2017, 55(2): 244-267.

[7] Chen Z, Zhai W, Wang K. Locomotive dynamic performance under traction/braking conditions considering effect of gear transmissions. Vehicle System Dynamics, 2018, 56(7): 1097-1117.

[8] Wang Z, Cheng Y, Allen P, et al. Analysis of vibration and temperature on the axle box bearing of a high-speed train. Vehicle System Dynamics, 2019: 1-24.

[9] Wang Z, Zhang W, Yin Z, et al. Effect of vehicle vibration environment of high-speed train on dynamic performance of axle box bearing. Vehicle System Dynamics, 2019, 57(4): 543-563.

[10] Li T, Sun W, Meng Z, et al. Dynamic investigation on railway vehicle considering the dynamic effect from the axle box bearings. Advances in Mechanical Engineering, 2019, 11(4): 1687814019840503.

[11] Li X, Jia L, Yang X. Fault diagnosis of train axle box bearing based on multifeature parameters. Discrete Dynamics in Nature and Society, 2015, 2015.

[12] Chen Z, Zhou Z, Zhai W, et al. Improved analytical calculation model of spur gear mesh excitations with tooth profile deviations. Mechanism and Machine Theory, 2020, 149: 103838.

[13] Liu J, Shao Y, Lim T C. Vibration analysis of ball bearings with a localized defect applying piecewise response function. Mechanism and Machine Theory, 2012, 56: 156-169.

[14] Chen DX, Mechanical Design Handbook, Part 1. Beijing: China Machine Press,2007.

[15] Harris T A, Kotzalas M N. Essential concepts of bearing technology. CRC press, 2006.

[16] Zhai W. Two simple fast integration methods for large-scale dynamic problems in engineering. International Journal for Numerical Methods in Engineering, 1996, 39(24): 4199-4214.

[17] Dormand J R, Prince P J. A family of embedded Runge-Kutta formulae. Journal of Computational and Applied Mathematics, 1980, 6(1): 19-26

Buckling and Postbuckling Analysis Method of Stretchable and Flexible Sensor Networks Based on ABAQUS

Shuguang Hu
State Key Laboratory of Mechanical Structures
Nanjing University of Aeronautics and Astronautics
Nanjing, China
sg-hu@nuaa.edu.cn

Yu Wang
State Key Laboratory of Mechanical Structures
Nanjing University of Aeronautics and Astronautics
Nanjing, China
wngyu@nuaa.edu.cn

Lei Qiu*
State Key Laboratory of Mechanical Structures
Nanjing University of Aeronautics and Astronautics
Nanjing, China
lei.qiu@nuaa.edu.cn

Abstract—**Aircraft smart skin technology requires the integration of large-scale and lightweight sensor networks in aircraft structures. To meet these requirements, the design of island-bridge structure is adopted, so that the sensor network can be manufactured in a limited scale and expanded in a large scale. Theoretical analysis, finite element simulation and experimental verification have been carried out on the mechanical properties of island-bridge structures. However, theoretical analysis is difficult to analyze irregular shaped structures, and the existing simulation researches mainly focus on the island-bridge structures with one unit and lack the analysis of networks. In this work, the finite element simulation method of island-bridge structure networks based on ABAQUS has been proposed, which includes buckling mode analysis and postbuckling deformation analysis. The initial buckling deformation state of the structure is extracted by the buckling mode analysis, and the postbuckling deformation state of the structure can be obtained by the postbuckling deformation analysis. The simulation method is applied on the analysis of the mechanical properties and deformation patterns of different island-bridge structure networks, including serpentine, fractal and irregular island-bridge structure networks. This shows that the proposed simulation method can well guide the design of stretchable sensor networks.**

Keywords—aircraft smart skin, structural health monitoring, sensor network, flexible and stretchable, island-bridge structure, simulation method

I. INTRODUCTION

Aircraft smart skin is an advanced technology for future aircrafts. Structural health monitoring (SHM) is one of the key functions of aircraft smart skin [1-2]. Among the existing SHM technology, the method of combining piezoelectric (PZT) sensor network and guided wave-based monitoring is popular [3-5]. However, for large-area aerospace structures, arranging sensors artificially is time-consuming and inefficient. Furthermore, there are also problems such as poor sensor consistency and additional weight caused by wires. The piezoelectric sensors and wires can be integrated manufacturing through flexible printed circuit (FPC) process, also known as SMART Layer [6-7], which can simplify wires layout, reduce network weight, and ensures the consistency of the PZTs. However, due to the limitation of manufacturing process, it is very difficult to directly manufacture large-area sensor networks.

In recent years, flexible and stretchable electronic technology has been widely researched and applied in medical, communication and entertainment fields, such as electronic skin [8], pen shaped flexible display [9] and so on. Based on flexible electronic technology, the sensor network can be made flexible and stretchable, and the applications of "small-scale manufacturing, large-scale expansion" can be realized, which is conducive to the arrangement of sensor networks on large-area aviation structures.

In order to make sensor network not only have good flexibility and stretchability, but also have the ability to transmit signal, many scholars have studied the island-bridge structure. In the island-bridge structure, the island is used to deploy sensor, while the bridge is used to arrange the signal transmission conductor. In recent years, researches on island-bridge structure have mainly focused on the serpentine and fractal island-bridge structures with good stretchability, the interconnects of which are serpentine lines and fractal lines. Zhang *et al.* [10-11] proposed a mechanical model for studying serpentine and self-similar fractal interconnects, and the relationship between structural stretchability and geometric parameters was analyzed through the model. Based on curved beam theory and plane-strain assumption, Widlund *et al.* [12] research the relationship between serpentine geometry and its mechanical properties. In addition to the theoretical analysis, many scholars also study the island-bridge structure by finite element simulation. Zhang *et al.* [13-14] designed stretchable batteries based on the fractal island-bridge structure and proposed a simulation method for the fractal island-bridge structure, which is faster in computation than the traditional method. Based on ABAQUS, Wang *et al.* [15-16] proposed the simulation methods for serpentine and fractal island-bridge structures and analyzed the designed large-scale stretchable sensor network. Similarly, Guo and Kim *et al.* [17] analyzed the effect of sensor network design on structural strain reduction through the simulation of tensile deformation of one-unit island-bridge structures. However, there are two main defects in previous researches. First, the theoretical model is difficult to analyze the stress, strain, and deformation of arbitrarily irregularly shaped island-bridge structures. Second, the existing simulation methods focused on the island-bridge structures with only one unit and cannot adequately account for the deformation of the structural network.

In this article, aiming at the defects in the previous researches, the simulation method of island-bridge structure networks based on ABAQUS is proposed, which can guide the design and development of stretchable sensor networks. Based on this simulation method, the stress, strain distribution, stretchability and deformation state of the island-bridge structure networks during stretching process are obtained and analyzed, including 1×3 array serpentine island-bridge structure, 2×3 array serpentine island-bridge structure, 2×2 array fractal island-bridge structure and 2×2 array irregular island-bridge structure.

---

*Lei Qiu is the corresponding author. (e-mail: lei.qiu@nuaa.edu.cn).

II. SIMULATION METHOD OF ISLAND-BRIDGE STRUCTURE NETWORKS

A. Simulation Model and Method

Due to the good stretchability of the serpentine and fractal interconnects, this article mainly focuses on the simulation method of the serpentine and fractal island-bridge structure networks. According to the array arrangement of islands, the island-bridge structure network can be described, which mainly includes: 1×3 array serpentine island-bridge structure, 2×3 array serpentine island-bridge structure, 2×2 array fractal island-bridge structure and 2×2 array irregular island-bridge structure. The serpentine and fractal geometries are shown in Fig. 1. The serpentine geometry is determined by the following five geometric parameters: width w, thickness t, number of units m, semi-circular arc diameter l_1, linear segment l_2. The fractal interconnect is a second-order serpentine-based fractal interconnect, the overall size of the structure is controlled by the height $h^{(2)}$ and spacing $l^{(2)}$, the local size is controlled by the serpentine interconnect.

During buckling deformation, the serpentine and fractal interconnects will undergo not only large deformation and displacement, but also large rotational deformation, which is a geometrically nonlinear problem. For the above problem, ABAQUS finite element analysis software has good simulation efficiency and accuracy. The simulation method of the island-bridge structure networks based on ABAQUS mainly includes two aspects, buckling mode analysis and postbuckling deformation analysis. When the island-bridge structure network is subjected to in-plane tensile load, the interconnects are prone to out-of-plane buckling and exist different buckling modes because of its large width-to-thickness ratio, potential material defects and external disturbances. Based on buckling mode analysis, different buckling modes of interconnects can be extracted, which reflect the possible deformation trend of interconnects. Further, buckling modes are introduced as an initial imperfection in postbuckling deformation analysis to ensure the accuracy of the simulation results. After the buckling mode is introduced, the deformation state of the structure and the structural stress, strain, and displacement distribution can be obtained by postbuckling deformation analysis. The simulation process is shown in Fig. 2.
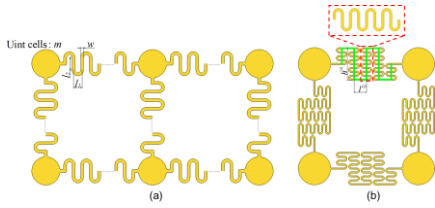
Fig. 1. Schematic of interconnects of island-bridge structure networks. (a) Serpentine interconnect. (b) Fractal interconnect.
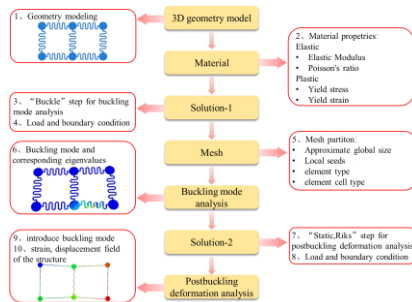
Fig. 2. Process of the simulation method.

1) 3D geometric model and material properties: The 3D geometric model can be created in the "Parts" module of ABAQUS. Since FPC is usually used to produce sensor network and polyimide is often used as the base material, polyimide (PI) is used in this simulation. The elastic strain limit of PI is 2%~3%, which is considered as 2% in this article. The simulation only needs to obtain the deformation state of the structure at the elastic strain limit of PI, in order to simplify the simulation, the material can be regarded as an ideal elastic-plastic material. In addition, when setting material properties, the unity of calculation in ABAQUS is important for simulation. The details are as follows: elastic modulus 2500 MPa, Poisson's ratio 0.34, yield stress and corresponding yield strain are 120 MPa and 0, respectively.

2) Solution setting: The solution setting includes the step, load and boundary condition. Furthermore, the solution setting for buckling mode analysis and postbuckling deformation analysis are different and will be further discussed below.

3) Mesh partition: The element shape is set as "HEX" and the element type is eight-node 3D reduced integration, hourglass controlled soli elements (C3D8R). When setting the global seeds, the global approximate size is set as 0.1 and the maximum deviation factor is 0.01, so that the element size is moderate and the distribution is reasonable. For simple serpentine interconnect, the simulation speed is faster, while for complex fractal interconnect, the simulation can be easy to converge. In addition, the local seeds must be performed on islands, which have smaller deformation, to increase the speed of simulation. It is worth noting that after the mesh partition, a mesh check is required. The check shows that there is no errors and warnings and then the simulation can be carried out.

B. Buckling Mode Analysis

The step of buckling mode analysis is "Buckle" step, and the buckling mode and corresponding critical load (the product of the eigenvalue and the concentrated force) can be obtained through buckling mode analysis. In the step setting, the number of eigenvalues requested should be as large as possible to ensure that the desired buckling mode can be obtained. During buckling mode analysis, it is also necessary to output the displacement values of each buckling mode for postbuckling deformation analysis.

In order to ensure that the simulation of the structure is consistent with the real stretching situation, the islands at the two ends or diagonal corners of the island-bridge structure network should be constrained. One of the islands is completely fixed, the other islands are stretched on the same horizontal plane, and the remaining islands are free. Concentrated force is set on the far right or bottom of the island. Moreover, in order to prevent local deformation from being too large, causing non-convergence, it is necessary to couple the point of concentrated force with the island through equation constraint.

C. Postbuckling Deformation Analysis

The step of postbuckling deformation analysis is "Static, Riks" step. Before the simulation, for giving an initial imperfection to structure in model copy, buckling modes corresponding to the smallest positive eigenvalue of each

independent interconnect need to be introduced respectively, which is the earliest buckling deformation state in the uniaxial stretching process. The way of introducing the displacement of buckling mode is to add the sentence "*imperfection, file=buckle, step=1(new line) 9,0.05" after sentence "*End Assembly". Among them, "buckle" is the job name of the buckling mode analysis, and "9,0.05" means that the 5% displacement values of the buckling mode 9 are introduced as the initial imperfection.

The postbuckling deformation of the structure no longer meets the small deformation assumption, so it is necessary to turn on "NIgeom" in the step setting. The maximum load scale factor in the termination standard is set to 1, and the maximum number of increments should be set as large as possible to ensure that the structure deformation is enough large. In addition, the initial arc length increment should be less than the critical load, and the minimum and maximum arc length increments should be set reasonably to ensure that the arc length increment is within its range during simulation process. For postbuckling deformation analysis, the boundary condition settings remain unchanged, but the magnitude of the load should be adjusted according to the actual situation of the simulation.

If there are many negative increment steps and even non-convergence in the early stage of the simulation, and the structure cannot be fully buckled and unfolded, the simulation can proceed smoothly through appropriately reducing the magnitude of the load or increasing the initial imperfection. Besides, if the error message "A zero displacement solution was found in the first iteration of a riks step" appears in the simulation, larger deformation of the structure can be obtained by adjusting mesh size and load.

III. SIMULATION ANALYSIS OF SERPENTINE ISLAND-BRIDGE STRUCTURE NETWORKS

A. 1×3 Array Serpentine Island-Bridge Structure

In this simulation, the thickness t is 0.05 mm, the width w is 0.4 mm, the number of units m is 4, the semicircular arc diameter l_1 is 1.2 mm, and the straight section l_2 is 1.8 mm. Furthermore, the island radius is 2 mm, and the radius at the junction between serpentine and island is 0.1 mm. The center distance (30 mm) of the island can be regarded as the initial length of the two ends of the serpentine interconnect. It is assumed that the overall tensile strain is the ratio of the elongation of the structure to the initial length, which indicates the degree of tensile deformation of the structure. When the structure strain reaches the maximum elastic strain, the overall tensile strain is considered as the stretchability of structure.

The 1×3 array serpentine island-bridge structure has symmetric and antisymmetric forms, and it can be explained that the interconnect is symmetric and antisymmetric relative to the middle island, as shown in Fig. 3. The middle island can be regarded as a part of the interconnect, so the 1×3 serpentine island-bridge structure has only one independent interconnect.

For the buckling mode analysis, the number of eigenvalues requested is set as 20, the concentrated force is 1 N, and the direction of the force is the tensile direction. For the postbuckling deformation analysis, 5% displacement values of the buckling mode are introduced as the initial imperfection, and the concentrated force is 0.1 N. In addition, in the step setting, the maximum increment step is 1000, the initial arc length increment is 1E-5, and the minimum arc length

increment is 1E-25, the maximum arc length remains unchanged.
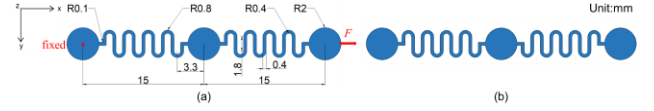
Fig. 3. Geometry schematic of 1×3 array serpentine island-bridge structure. (a) Antisymmetric 1×3 array serpentine island-bridge structure. (b) Symmetric 1×3 array serpentine island-bridge structure.

The displacement and strain distribution of the structure can be obtained from the simulation, which can be used to study the deformation state and stretchability of the structure. The stretchability of symmetric and antisymmetric structures are basically the same, as shown in Fig. 4 and Fig. 5. When the maximum strain of the structure reaches about 2%, the overall tensile strain is 90%, and the location of the maximum strain is at the junction between the island and the serpentine. Because the radius of arc is too small, stress concentration occurs at the junction, which leads to greater strain. In fact, the strain can be reduced by increasing the radius of arc, so the area of junction can be hidden for analysis. After that, when the overall tensile strain is about 90%, the maximum strain of the structure is only 1.3%, which is located at the top of the arc in the serpentine. In addition, when the maximum strain of the structure is 2%, the overall tensile strain of the structure is about 95%, so the stretchability of structure can reach about 95%.

During the stretching, the rotation of the middle islands of the symmetric and antisymmetric structure network is different. As shown in Fig. 6, the maximum rotation angle of the middle island of the antisymmetric structure can reach more than 10°, while the middle island of the symmetric structure is very small. Therefore, the antisymmetric structure isn't conducive to orientation-sensitive sensors, such as strain sensors.
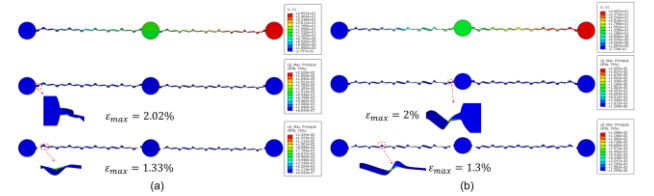
Fig. 4. Displacement and strain distribution of 1×3 array serpentine island-bridge structure. (a) Antisymmetric 1×3 array serpentine island-bridge structure. (b) Symmetric 1×3 array serpentine island-bridge structure.
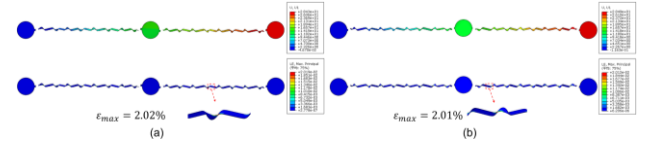
Fig. 5. Displacement and strain distribution of 1×3 array serpentine island-bridge structure after hiding junction. (a) Antisymmetric 1×3 array serpentine island-bridge structure. (b) Symmetric 1×3 array serpentine island-bridge structure.
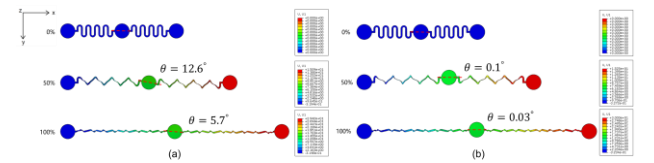
Fig. 6. The rotation angle of the middle island of 1×3 array serpentine island-bridge structure. (a) Antisymmetric 1×3 array serpentine island-bridge structure. (b) Symmetric 1×3 array serpentine island-bridge structure.

B. 2×3 Array Serpentine Island-Bridge Structure

In the 2×3 array serpentine island-bridge structure, the geometric parameters of the serpentine interconnect are the same as the above, and there are also symmetric and antisymmetric forms, as shown in Fig. 7. Initially the 2×3 array of serpentine island-bridge structure covers a 30 mm×15 mm rectangular area, which regards as a reference to measure the degree of expansion of the structure.

The middle islands are connected by a serpentine interconnect, due to deformation coordination, the four serpentine interconnects at both ends of the middle island and the interconnect between the middle islands will influence each other in structure deformation. The structure can be considered as having three independent interconnects.

For the buckling mode analysis, the number of eigenvalues requested is set as 40, and the concentrated force is set as 1 N. For the postbuckling deformation analysis, 5% and 10% displacement values of the buckling mode, corresponding to symmetric and antisymmetric structure respectively, are introduced as the initial imperfection. The force is set as 0.1 N. In addition, in the step setting, the initial arc length increment is set to 1E-5, and the other settings are the same as the 1×3 array serpentine island-bridge structure.

Similarly, the stretchability of the antisymmetric and symmetric structures are basically the same, as shown in Fig. 8. When the structure is expanded to the size of 360% of its initial area, the maximum strain of the structure is about 2%, and the location of the maximum strain is at the junction of the serpentine interconnect and the island. Besides, after hiding the junction, the maximum strain of the structure is reduced to 1.3%, and the maximum strain location is also moved to the top of the arc in the serpentine.

During the stretching process, the rotation of the middle island of the symmetric and antisymmetric structure is also different. As shown in Fig. 9, during stretching, the maximum rotation angle of the middle island of the symmetric structure is only 1°, while the maximum rotation angle of the middle island of the antisymmetric structure reaches about 15°. Similarly, if an orientation-sensitive sensor is placed on the antisymmetric structure, the measurement result will be greatly affected.
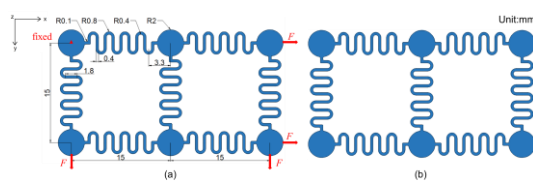
Fig. 7. Geometry schematic of 2×3 array serpentine island-bridge structure. (a) Antisymmetric 2×3 array serpentine island-bridge structure. (b) Symmetric 2×3 array serpentine island-bridge structure.
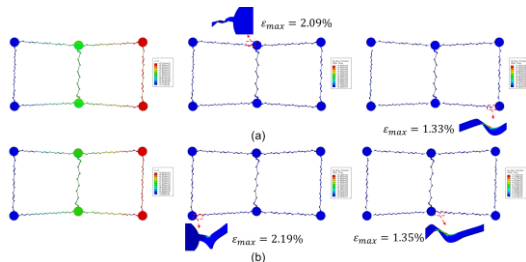
Fig. 8. Displacement and strain distribution of 2×3 array serpentine island-bridge structure. (a) Antisymmetric 2×3 array serpentine island-bridge structure. (b) Symmetric 2×3 array serpentine island-bridge structure.
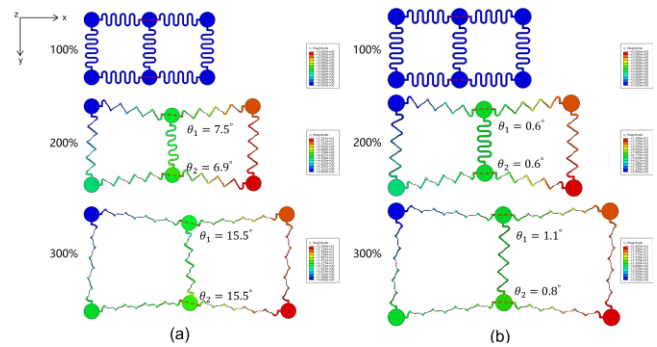
Fig. 9. The rotation angle of the middle island of 2×3 array serpentine island-bridge structure. (a) Antisymmetric 2×3 array serpentine island-bridge structure. (b) Symmetric 2×3 array serpentine island-bridge structure.

IV. SIMULATION ANALYSIS OF FRACTAL ISLAND-BRIDGE STRUCTURE NETWORKS

A. 2×2 Array Fractalr Island-Bridge Structure

The interconnects are serpentine-based second-order fractal interconnects. As shown in Fig. 10, from the whole, the fractal interconnect can be viewed as serpentine interconnect with only one unit, and the local straight segment is composed of the serpentine interconnect with four units, the size of the local serpentine interconnect has been described above. The initial area enclosed by the structure, which is 24.4 mm×24.4 mm, can be regarded as a reference to describe the degree of expansion of the structure.

The four islands are only stretched on the same horizontal level, and the deformation of interconnects is only affected by the degree of tension between the two islands. Therefore, the 2×2 array fractal island-bridge structure has four independent interconnects.

In the buckle step setting, the number of eigenvalues requested is set to 40. In the postbuckling deformation step setting, the initial arc length increment is set to 1E-6, and the rest is the same as above. Before postbuckling deformation analysis, 50% displacement values of the buckling mode must be introduced as the initial imperfection. The force is set as 1 N in the buckling mode analysis, and 0.02 N in the postbuckling deformation analysis.
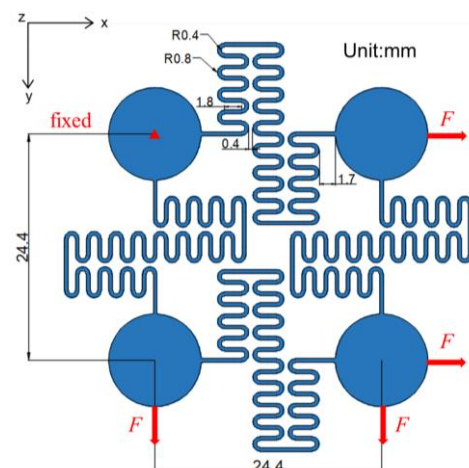
Fig. 10. Geometry schematic of 2×2 array fractal island-bridge structure.

The simulation results are shown in Fig. 11. When the structure is stretched to 2500% of the initial area, the maximum strain is 0.65%, which is located at the top of inner arc of the interconnect. It can be seen that the stretchability of

the fractal interconnect is much greater than the serpentine, and the position of the maximum strain is at the top of the inner arc.
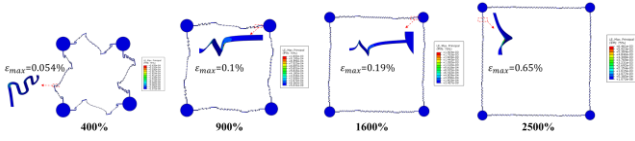
Fig. 11. Strain distribution of 2×2 array fractal island-bridge structure.

Due to the different sequences of appearing buckling critical loads, the sequence of appearing the buckling deformation is different. Furthermore, after appearing the buckling deformation, the strain energy of the structure will decrease. Therefore, under the same force, deformation of the interconnect which first appears buckling deformation is large, and the deformation of other interconnects are smaller, resulting in the phenomenon that the structure does not expand in a square. In the postbuckling deformation analysis, the structure can expand in a square shape by changing the concentrated force to the displacement load.

B. 2×2 Array Irregular Island-Bridge Structure

It can be seen that the fractal interconnects have a high utilization rate of the blank area enclosed by structure. Even so, the length l_2 of the local serpentine interconnect segment still can be increased to optimize the structure and further use the blank area, as shown in Fig. 12. However, the optimized fractal interconnects are irregular, and the simulation analysis is more complicated.

Due to the complexity of the optimized structure, if concentrated force is used, the simulation is prone to non-convergence. Therefore, the simulation adopts displacement load. Similarly, in the buckling mode analysis, the number of eigenvalues requested is set to 40. In the postbuckling deformation analysis, the maximum increment step is 2000, the initial arc length increment is 1E-7, the minimum arc length increment is 1E-25, and the maximum arc length increment is unchanged. In addition, 50% displacement values of the buckling mode should be introduced as the initial imperfection.

The simulation results are shown in Fig. 13. When the structure is stretched to 2500% of the initial area, the maximum strain is 0.091%. Similarly, the maximum strain is also located at the top of the inner arc. After optimizing, the internal strain level is greatly reduced and the stretchability is enhanced.
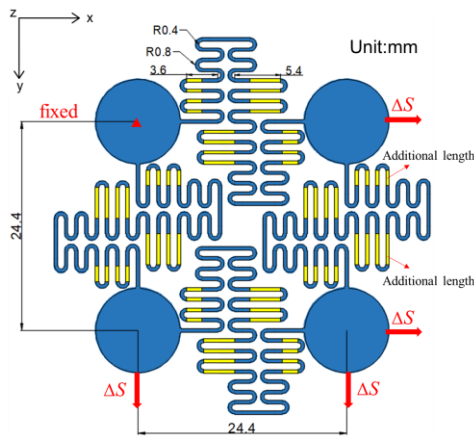
Fig. 12. Geometry schematic of 2×2 array irregular island-bridge structure.
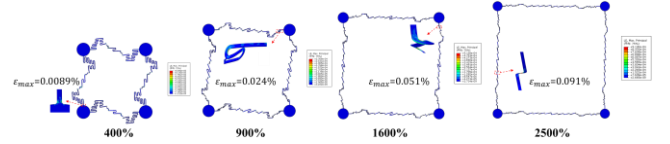
Fig. 13. Strain distribution of 2×2 array irregular island-bridge structure.

V. CONCLUSION

In this article, the simulation method of island-interconnect structure networks based on ABAQUS is proposed, which can guide the design of stretchable sensor networks, including buckling mode analysis and postbuckling deformation analysis. Based on this simulation method, the buckling deformation and mechanical properties of different type island-bridge structure networks can be analyzed, including serpentine, fractal and irregular island-bridge structure networks. The simulation results show that the stretchability of the 1×3 array serpentine island-bridge structure is 95% and the 2×3 array serpentine island-bridge structure can be stretched to 360% of its initial area. However, the maximum rotation angle of the middle island of the symmetric structure is only 1°, while that of the antisymmetric structure is larger than 10°. For orientation-sensitive sensors, they cannot be arranged on the antisymmetric structure. Besides, for the 2×2 array fractal island-bridge structure, when stretched to 2500% of its initial area, the maximum strain of the structure is 0.65%, while for the optimized irregular structure, the maximum strain of the structure is reduced to 0.091%, the stretchability is greater enhanced. In conclusion, the simulation method is simple and suitable for buckling deformation analysis of different type island-bridge structure networks to further guide and optimize the design of large-scale sensor networks. However, as the complexity of the structure increases, the amount of simulation calculation will greatly increase. The simulation method needs to be further optimized to reduce the amount of calculation.

REFERENCE

[1] P. D. Foote, "Integration of structural health monitoring sensors with aerospace, composite materials and structures," Mat.-wiss. u. Werkstofftech, vol. 46, pp. 197-203, February 2015.

[2] L. Qiu, X. Deng, S. Yuan, Y. Huang, and Y. Ren, "Impact monitoring for aircraft smart composite skins based on a lightweight sensor network and characteristic digital sequences," Sensors, vol. 18, pp. 2218, July 2018.

[3] V. Giurgiutiu, "Tuned Lamb wave excitation and detection with piezoelectric wafer active sensors for structural health monitoring," J. Intell. Mater. Syst. Struct., vol. 16, pp. 291-305, April 2005.

[4] L. Qiu, M. Liu, X. Qing, and S. Yuan, "A quantitative multidamage monitoring method for large-scale complex composite," Struct. Health Monit., vol. 12, pp. 183-196, March 2013.

[5] L. Qiu, B. Liu, S. Yuan, and Z. Su, "Impact imaging of aircraft composite structure based on a model-independent patial-wavenumber filter," Ultrasonics, vol. 64, pp. 10-24, January 2016.

[6] M. Lin and F. K. Chang, "The manufacture of composite structures with a built-in network of piezoceramics," Compos. Sci. Technol., vol. 62, pp. 919-939, June 2002.

[7] L. Qiu, F. Yuan, X. Shi and T. Huang, "Design of piezoelectric transducer layer with electromagnetic shielding and high connection reliability," Smart Mater. Struct., vol. 21, pp. 075032, June 2012.

[8] C. Pang, C. Lee and K. Y. Suh, "Recent Advances in Flexible Sensors for Wearable and Implantable Devices," J. Appl. Polym. Sci., vol. 130, pp. 1429-1441, August 2013.

[9] S. Forrest, "The path to ubiquitous and low-cost organic electronic appliances on plastic," Nature, vol. 428, pp. 911-918, April 2004.

[10] Y. Zhang, S. Xu, H. Fu, J. Lee, J. Su, and K. Hwang, "Buckling in serpentine microstructures and applications in elastomer-supported ultra-stretchable electronics with high areal coverage," Soft Matter, vol. 9, pp. 8062-8070, June 2013.

[11] Y. Zhang, H. Fu, Y. Su, S. Xu, H. Cheng, J. A. Fan, "Mechanics of ultra-stretchable self-similar serpentine interconnects," Acta Mater., vol. 61, pp. 7816-7827, December 2013.

[12] T. Widlund, S. Yang, Y. Y. Hsu, and N. Lu, "Stretchability and compliance of freestanding serpentine-shaped ribbons," Int. J. Solids Struct., vol. 51, pp. 4026-4037, November 2014.

[13] S. Xu, Y. Zhang, J. Cho, J. Lee, X. Huang, L. Jia, "Stretchable batteries with self-similar serpentine interconnects and integrated wireless recharging systems," Nat. Commun.,vol. 2013, pp. 1543, February 2013.

[14] Y. Zhang, H. Fu, S. Xu, J. A. Fan, K. C. Hwang, J. Jiang, "A hierarchical computational model for stretchable interconnects with fractal-inspired designs," J. Mech. Phys. Solids,vol. 72, pp. 115-130, December 2014.

[15] Y. Wang, Y. Luo, and L. Qiu, "Simulation Method of an Expandable Lamb Wave Sensor Network for Aircraft Smart Skin," IEEE Sens. J., vol. 20, pp. 102-112, January 2020.

[16] Y. Wang, L. Qiu, Y. Luo, and R. Ding, "A stretchable and large-scale guided wave sensor network for aircraft smart skin of structural health monitoring," Struct. Health Monit., June 2019.

[17] Z. Guo, K. Kim, N. Salowitz, G. Lanzara, Y. Wang, P. Yinan, "Functionalization of stretchable networks with sensors and switches for composite materials," Struct. Health Monit., vol. 17, pp. 598-623, May 2018.

Visual Inertial Calibration of Mobile Robotic System Based on Reinforcement Learning

Wenxing Zhu
Key laboratory of micro-inertial instrument and advanced navigation technology, Ministry of education
School of Instrument Science and Engineering
Southeast University
Nanjing,China
zwxseu@163.com

Lihui Wang*
Key laboratory of micro-inertial instrument and advanced navigation technology, Ministry of education
School of Instrument Science and Engineering
Southeast University
Nanjing,China
wlhseu@163.com

Abstract—**Visual inertial calibration is an important part of Visual-Inertial System (VINS), and inaccurate visual inertia parameters will lead to unreliable navigation attitude estimation. Aiming at the problems of the complicated offline calibration process and the large amount of online self-calibration calculation, a novel visual inertial self-calibration method based on reinforcement learning (RL-VIC) is proposed to provide high precision calibration parameters for the lifetime autonomous operation. Firstly, the nonlinear observable analysis of the visual inertial system is carried out. The rank decomposition of the Fisher Information Matrix（FIM）is adopted to determine the observable parameter space. Then, reinforcement learning methods are used to determine unobservable discrete motion sequences, give different reward feedback to motion sequences with different parameters and different observability, and learning the best calibrated motion sequences. Finally, experiments demonstrate that the visual inertial calibration based on reinforcement learning method has a translation error of 0.0349m and a rotation error of 0.399°. Compared with the traditional offline calibration method Kalibr and the online automatic self-calibration method VINS-Mono, which uses a complete motion sequence for calibration, it greatly saves the calculation cost and obtains the calibration parameters with the same accuracy.**

Keywords—Visual inertial calibration, observability analysis, reinforcement learning strategies.

I. INTRODUCTION

Visual-Inertial Simultaneous Location and Mapping (VI-SLAM) has become an increasingly popular method in mobile robotic system. Based on the fusion of visual and inertial navigation information, VI-SLAM is aimed at tracking the position and posture of the subject during the movement, and building a consistent map of environmental structure at the same time[1]. The camera works well in most texture-rich scenes, but basically does not work in scenes with few features, such as white walls. In the meanwhile, the navigation error of inertial measurement units (IMU) accumulates over time, but it has high navigation accuracy in a short time. Complementing with each other, the camera and IMU can provide accurate state estimation and environmental structure with scale information. Because of the small size, low power consumption and cheap price, they are being deployed into robot navigation[2], unmanned aerial vehicles [3], three dimensional reconstruction[4], and augmented reality [5].

However, all VI-SLAM algorithms require camera-IMU external calibration parameters. Precise external parameters are required in obtaining the camera pose and performing the backend state optimization of camera and IMU data. Inaccurate external calibration parameters will cause systematic errors in motion estimation, reducing the overall performance of the system, and even tracking failure.

Calibration parameters, such as camera intrinsic parameters, scale factors, external parameters rotation and translation, and etc. are common to all visual inertial mobile robotic systems, which are essential for state estimation, path planning and control. Usually, calibration parameters are obtained by specific and complex calibrations routines offline, or a self-calibration framework relies on sufficiently stimulated motion [6]. In the complicated and time-consuming process of offline calibration, the technicians are required to move the calibration checkboard in front of the device carefully. In addition, when the external parameters of camera-IMU are changed due to the collision or adjustment of the sensor, the offline calibration process needs to be repeated regularly to calibrate the device. Similarly, visual inertial calibration based on online self-calibration is highly dependent on accurate initialization parameters. The online self-calibration will greatly increase the dimension of the state space without providing other measurement values. And the task of collecting measurement values which make the entire state space observable is not easy. Meanwhile, more complicated initialization process and higher calculation cost is required in automatic self-calibration. In the case of life-term operation, visual inertial self-calibration systems are necessary because they can compensate for calibration errors caused over time, such as a collision or sensor replacement [7]. Self-calibration has attracted great attention in mobile robotic system, and the slow time-varying drift has proven to be a challenge.

To solve the existing calibration problem, a novel visual inertial self-calibration method is proposed based on reinforcement learning. Observability is analyzed by using the connection between FIM and nonlinear observability. The singular threshold of the FIM is used to detect the directions in the unobservable parameter space, and avoiding updating the state estimates in the directions corresponding to the unobservable parameter space. Reinforcement learning methods are used to determine unobservable discrete motion sequences, give different reward feedback to motion sequences with different parameters and different observability, and learning the best calibrated motion sequences. The action space is discretized into the actions easily performed by the calibrator, and the state space is defined by parameter states. Our inspiration is to obtain

---

different rewards through observability analysis of the dataset sequence, and continuous interaction with the data set motion sequence through reinforcement learning.

II. RELATED WORK

Visual-inertial calibration has been studied extensively and many algorithms have sprung up. Camera-IMU parameter calibration mainly includes two methods: offline calibration and automatic online self-calibration.

A. Offline Calibration

Rehder J. proposed the camera-IMU external parameter offline calibration tool Kalibr for the first time, and improved it in the subsequent work[8]. Considering the coupling of calibration error of multiple IMUs' axis, then the pixel photometric value is used to directly calibrate the external parameters[9]. These work greatly improved the accuracy of the camera-IMU external parameter offline calibration, and providing reliable camera-IMU external calibration parameters for the VI-SLAM systems. However, in the complicated and time-consuming process of offline calibration, the technicians are required to move the calibration checkboard in front of the device carefully[10]. In addition, when the external parameters of camera-IMU are changed due to the collision or adjustment of the sensor, the offline calibration needs to repeat the calibration process to re-calibrate the parameters.

B. Automatic Online Self-calibration

Automatic online self-calibration means calibrating the camera-IMU parameters online without any auxiliary equipment and calibration technicians. The automatic online calibration method brings convenience to the rapid calibration, while more complicated initialization process and higher calculation cost is required.

Extended Kalman Filter (EKF) is proposed by Li M and Mourikis A I. to estimate the state vector of monocular camera-IMU external parameter[11], but the algorithmic convergence highly depends on the accuracy of initial state estimation. Meanwhile, they introduced an online algorithm for calibrating the time offset between the camera and the IMU in an EKF based visually assisted inertial navigation system [12]. In addition, Kelly J improved an unscented Kalman filter (UKF) for self-calibration and relative attitude [13]. Mirzaei F M also proposed an EKF for calibrating the unknown transformation between the camera and the IMU accurately. In order to interpret the observability of the nonlinear system in the visual inertial calibration process, observability grade standard based on Li derivative is adopted in this article [14]. Tang creatively introduced the constrained Gaussian filters to calibrate all the external parameters online for the visual odometer navigation system [15]. This work has great significance for the camera-IMU external parameters calibration. Yang Z. proposed a novel method that could initialize the speed, gravity, scale, and calibrate camera-IMU transformation accurately [16]. Compared with the latest marked offline calibration method, the calibration results show the excellent performance with 0.02 m of the translation errors and 1° of the rotation errors. Observability analysis of IMU camera calibration parameters is essential. By using the Lie derivative rank condition, Panahandeh G. derived the nonlinear state space model and analyzed the observability [17].

Both offline calibration and automatic self-calibration have their own shortcomings and algorithm flaws. To solve the existing calibration problem, a novel visual inertial self-calibration method is proposed based on reinforcement learning, which can save computation space and does not require manual calibration operations.

III. METHODOLOGY

The proposed algorithm has three parts: VIO parameter initialization, nonlinear observable performance analysis, and self-calibration based on reinforcement learning. The structure of the RL-VIC is shown in Fig.1. Firstly, the data information of the camera and IMU is initialized by visual inertial odometer. Secondly, the FIM matrix is used to establish the relationship between the parameters to be calibrated and the non-linear observability. Then, a reinforcement learning method is proposed to train the visual inertial data set. Through the FIM observability analysis, which motion sequences can be subjected to complete parameter space regression, and reward feedback is given to the sequence, and the best calibration data is selected from the entire data for calibration.
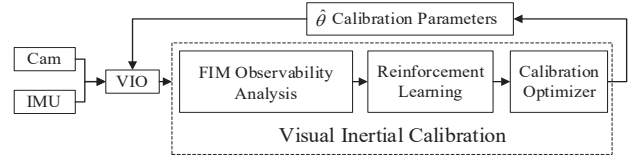
Fig. 1. The structure of the RL-VIC

A. Visual Inertial Self-calibration

The widely used camera model is the pinhole model, which is given by the center projection transformation from a three-dimensional space to a plane.

The spatial point $M^c = (x_c, y_c, z_c)^T$, and the corresponding image point is $m = (x, y)^T$. According to the triangle similarity principle, it can be inferred that the spatial point M and its image point m satisfy the following relationship

$$\begin{cases} x = f x_c / z_c \\ y = f y_c / z_c \end{cases} \quad (1)$$

$M^w = [x_w, y_w, z_w, 1]^T$ represents the homogeneous form of point M in the world system. Then the coordinate transformation relationship between M^c and M^w can be expressed as follows

$$M^c = \begin{vmatrix} R_{cw} & t^c_{cw} \\ 0_{1\times3} & 1 \end{vmatrix} M^w \quad (2)$$

The ideal camera projection model can be derived as follows

$$P = z_c \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_o & 0 \\ 0 & f_v & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{cw} & t^c_{cw} \\ 0_{1\times3} & 1 \end{bmatrix} = K \begin{bmatrix} R_{cw} | t^c_{cw} \end{bmatrix} \quad (3)$$

where $f_u = f / dx$, $f_v = f / dy$ are the horizontal and vertical focal lengths of the camera expressed in pixel units respectively. $\begin{bmatrix} R_{cw} | t^c_{cw} \end{bmatrix}$ the external camera parameters. Focal length and principal point coordinates become the camera internal parameters, expressed by matrix K

$$K = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

The low-cost MEMS is considered. Assuming there is only the rotation between the two sensors, and the translation between the accelerometer and the gyroscope is negligible. The model for the gyroscope measurements $\tilde{\omega}$ is given by

$$\tilde{\omega} = \omega_{GI} + b_g + \eta_g \tag{5}$$

where ω_{GI} represents the projection of the angular velocity in the inertial coordinate system. The random walk process b_g is given by

$$\dot{b}_g = \eta_{bg} \tag{6}$$

Zero mean Gaussian random white noise η_g and η_{bg} is defined as

$$\eta_g \sim N\left(0, \sigma_g^2 \cdot I_3\right) \tag{7}$$

$$\eta_{bg} \sim N\left(0, \sigma_{bg}^2 \cdot I_3\right) \tag{8}$$

Similarly, accelerometer output can be modeled as

$$\tilde{a} = R_{WI}^T \cdot (a_{WI}^w - g^w) + b_a + \eta_a \tag{9}$$

where a_{WI}^w the projection of carrier acceleration in world system. R_{WI}^T the corresponding direction cosine matrix. b_a the bias vector can be defined as a random walk process.

Camera measurement can be interpreted as the projection of the linear velocity and angular velocity under the camera system. To establish an error term and related objective function, the difference between the actual sensor measurement and the predicted value based on the current state is calculated.

The accelerometer error is

$$e_{bk} = \tilde{b}_k^I - R_{WI}(\varphi(t_k))^T (a_{WI}^w(t_k) - g^w) - b_a \tag{10}$$

The gyroscope error is

$$e_{\omega k} = \tilde{\omega}_{WI,k}^I - \xi_{WI}^I(t_k) - b_g \tag{11}$$

Camera speed error is

$$e_{ck} = \tilde{v}_{WC,k}^C - R_{IC}^T [R_{WI}(\varphi(t_k)^T) v_{WI}^W(t) + \omega_{WI}^I \times t_{IC}] \tag{12}$$

Camera angular velocity error is

$$e_{wc,k} = \tilde{\omega}_{WC,k}^C - R_{IC}^T \omega_{WI}^I \tag{13}$$

The final optimization objective function consists of four items

$$\begin{aligned} F &= F_b + F_\omega + F_v + F_{\omega c} \\ &= \frac{1}{2}\sum_{k=1}^{K} e_{bk}^T R_{bk}^{-1} e_{bk} + \frac{1}{2}\sum_{k=1}^{K} e_{\omega k}^T R_{\omega k}^{-1} e_{\omega k} + \frac{1}{2}\sum_{k=1}^{K} e_{vk}^T R_{vk}^{-1} e_{vk} + \frac{1}{2}\sum_{k=1}^{K} e_{wc,k}^T R_{wc,k}^{-1} e_{wc,k} \end{aligned} \tag{14}$$

Parameter vector to be optimized Θ

$$\Theta = \left\{\varphi_{IC}^T, t_{IC}^T, b_a^T, b_g^T, g^{IT}\right\} \tag{15}$$

where φ_{IC}^T, t_{IC}^T the rotation vector and translation vector between IMU and camera.

B. Observability of Visual Inertial Self-calibration

Jauffret C.[18]established the connection between the reversibility of FIM and the observability of parameter estimation in nonlinear regression problems. Visual inertial calibration can be solved only when the FIM is reversible. Given the current observations, the singular rank of FIM corresponds to unobservable directions in calibration parameter space. There are some parameters can be observed, however, there is no guarantee that all parameters can be observed in the collected calibration dataset. By performing singular value decomposition (SVD) on FIM matrix and analyzing singular values and numerical observability, matrix with insufficient numerical rank could be identified.

To distinguish the situation the same scale matrix entries affected by noise interference, the numerical rank is carried out to determine the singular values. The scaling matrix S can be computed as

$$S = \text{diag}\left\{\frac{1}{\|F(:,1)\|}, \cdots, \frac{1}{\|F(:,n)\|}\right\} \tag{16}$$

where $\|F(:,n)\|$ the column norm of the Jacobian matrix. By using Cholesky decomposition, the error covariance matrix G could be decomposed into square root form. The rank revealing decomposition are used to estimate the numerical rank of the FIM directly. Dealing with SVD rank decomposition on matrix (LJ).

$$LJ = USV^T \tag{17}$$

where $S = \text{diag}(s_1, \ldots, s_n)$ represents the singular values.

FIM is based on the inverse Hessian approximation of Gauss-Newton approximation. The accuracy of certain parameters is determined by the likelihood function of the FIM. In this way, it can be regarded as the accuracy matrix of the Maximum Likelihood Estimation (MLE).

C. Visual Inertial Calibration Based on Reinforcement Learning

Without labeled training data in visual inertial dataset, there is no need to find the hidden structure in the data. Meanwhile, not all subsequences can regress the calibration parameter space. The unobservable motion sequence which is used to calibrate the calibration parameters will bring catastrophe. To solve the existing calibration problem, a novel visual inertial self-calibration method based on reinforcement learning (RL-VIC) is proposed. Our inspiration is to obtain different rewards through observability analysis of the dataset sequence, and continuous interaction with the data set motion sequence through reinforcement learning. Select the best reward sequence for visual inertial calibration, so that the calibration parameters obtained are accurate and save computing resources.

Not all motion sequences can obtain accurate calibration parameters. For an arbitrary carrier motion sequence, a large part of which will not make the calibration parameters linear regression. The usual calibration data collection takes a lot of time to collect enough action excitation, so that obtaining the linear regression of the entire calibration parameter space. Based on the previous nonlinear observability combined with reinforcement learning, the motion sequence which leads to the regression of the calibration space could be determined. The calibration problem can be interpreted as Partially Observable Markov Decision Process (POMDP). The action space is discretized into the actions easily performed by the calibrator, and the partially observable state space is defined by parameter states.

POMDP refers to a framework that learns from interaction to achieve goals. The environment response to the agent's behavior and gives the agent a reward function. The goal of the agent is to maximize the accumulated reward function. The agent and the environment interact in a discrete time series, receive the environment state information, and chooses actions. At time t+1, the agent receiving the value representing the reward brought by the selected action, and in a new state, the process of the agent interacting with the environment can be done by $S_0, A_0, R_1, S_1, A_1, R_2, S_2, \cdots$ sequence representation.

29

To maximize its own cumulative income, the agent has the precise representation of income G_t. It is defined as a certain specific function of the reward sequence starting from the t time step. In the simplest case, G_t is the cumulative sum of all rewards when the agent interacts with the environment, as shown in the following formula

$$G_t = R_{t+1} + R_{t+2} + ... + R_T \qquad (18)$$

where T is the last time step in the interaction process, and S_T is the termination state.

The value function is divided into state value and state-action pair value, denoted by v and q respectively. The state value function refers to the expected cumulative return of the agent from the state S under a certain strategy π.

$$v(s) = E[G_t \mid S_t = s] \qquad (19)$$

Markov reward process refers to Markov sequence with value judgment. Its form is a quaternion $\langle S, P, R, \gamma \rangle$, where γ is the discount factor, R is the reward function, and R_s is the expectation of the reward in the state

$$R_S = E[R_{t+1} \mid S_t = s] \qquad (20)$$

Expand G_t and decompose the value function recursively to get

$$v(s) = E[R_{t+1} + \gamma(R_{t+2} + \gamma R_{t+3} + ...) \mid S_t = s] \qquad (21)$$

Similarly, for the state action value function

$$q_\pi(s,a) = E[R_{t+1} + \gamma q_\pi(S_{t+2}, A_{t+1}) \mid S_t = s, A_t = a] \qquad (22)$$

The state space includes the calibration parameter state vector u_i, the action is control input, and the process model of robotic motion $X_{i+1} = f(X_i, u_i)$ will provide the next pose. Only one confidence state $\hat{X}$ can be estimate with the sufficient observability. The MDP formula defined follow

$$\begin{cases} s = \hat{\theta} \\ a = \{X_i, ..., X_{i+k}\} \\ r = +r, if \ cov(\hat{\theta}_j) < \varepsilon \end{cases} \qquad (23)$$

where $\hat{\theta}$ the estimated calibration parameter vector, $\{X_i, ..., X_{i+k}\}$ the sequence of actions performed on the robotic system. The reward $+r$ is adjustable related to the specific estimated parameters.

The transition model $P(s'|a,s)$ function gives the possibility that a new state s' will be obtained from the state s based on the action a. When the Q-value is correct, the optimal actions can be learned without knowing the model, and the following constraint equation to ensure balance is deduced as

$$Q_*(s,a) = E[R_{t+1} + \gamma \sum Q_*(S_{t+1}, a') \mid S_t = s, A_t = a]$$
$$= \sum_{s' \in S} P(s' \mid s, a)[r + \gamma \max_{a'} Q_*(s', a')] \qquad (24)$$

where formula (24) is the Bellman optimal equation Once the Q-value converges, the learned policy $\pi(s)$ from the current state is

$$\pi(a,s) = \arg\max Q_a(s,a)$$
$$= \max_a \sum_{s',r} p(s',r \mid s,a)[r + \gamma \pi(s')] \qquad (25)$$

Once knowing $Q_*(s,a)$, it is relatively easy to find the optimal strategy. After understanding the information of $Q_*(s,a)$ in any state, the agent needs to do a single-step search to find the state with the greatest value among the next reachable states, until the end stop state and get the optimal strategy.

In both internal and external situations, the reward table is simple: the reward 5 is given for intermediate state transitions, and the reward 20 is given for ultimate state transitions. The reward value is selected by manual adjustment. The method is available is because the number of states can be managed through discrete Q-learning. The parameters of the reinforcement learning are initialized as follows: Q-value is zero, and use $\gamma = 0.98$, $\alpha = 0.10$. The outer loop is executed to meet the convergence criteria, the upper limit of the number $iter = 200$ is set as the maximum value of the iteration.

IV. EXPERIMENTS

To Interpret the calibration performance of the novel visual inertial self-calibration method based on reinforcement learning, the EuRoc V1-01 easy dataset is used to test the algorithm in this paper. In the visual inertial dataset, not all sub-sequences can regress the calibration parameter space. The unobservable motion sequence is used to calibrate the calibration parameters that will bring catastrophe. The proposed algorithm RL-VIC is to obtain different rewards through observability analysis of the data set sequence, and continuous interaction with the data set motion sequence through reinforcement learning. Select the best reward sequence for visual inertial calibration, so that the calibration parameters obtained are accurate and save computing resources. but achieves the same accuracy of calibration parameters. The accuracy of the calibration algorithm is compared with the traditional offline calibration method Kalibr and the online automatic self-calibration method VINS-Mono.

TABLE I. demonstrates the comparison of calibration parameters based on reinforcement learning and traditional methods. f_X, f_Y represents the focal length of X axis and Y axis. C_u, C_v represents the principal point.

TABLE I. COMPARISON INTRINSIC CALIBRATION ERROR RESULTS BETWEEN REINFORCEMENT LEARNING AND TRADITIONAL METHODS

Parameters	Reinforcement Learning (best calibration sequence)	Traditional Kalibr (complete dataset)
$f_X[px]$	244.58±0.68	246.45±0.35
$f_Y[px]$	245.44±0.61	246.49±0.28
$C_u[px]$	322.67±0.57	324.45±0.45
$C_v[px]$	267.40±1.21	269.07±0.78

TABLE II. COMPARISON EXTRINSIC CALIBRATION ERROR RESULTS BETWEEN REINFORCEMENT LEARNING AND TRADITIONAL METHODS

Parameters	Reinforcement Learning	VINS Mono
yaw-error[degree]	-0.244	-
pitch-error[degree]	-0.281	-
roll-error[degree]	0.250	-
R_RMSE[degree]	0.399	0.26
x-error[m]	-0.036	-
y-error[m]	-0.040	-
z-error[m]	0.0278	-
T_RMSE[m]	0.0349	0.069

TABLE II. demonstrates the comparison of calibration parameters based on reinforcement learning and traditional methods. yaw-error, pitch-error, roll-error represents the

30

rotation error of yaw, pitch and roll. *x-error* , *y-error* , *z-error* represents the translation error of X, Y and Z. *R_RMSE* and *T_RMSE* represents the Root Mean Square Error (RMSE) of rotation and translation Respectively . From the experiment results, the rotational RMSE of proposed method RL-VIC is 0.399 degree, while the VINS-Mono is 0.26 degree. The translational RMSE of proposed method RL-VIC is 0.349 m, while the VINS-Mono is 0.069m.

Fig.2 demonstrates the velocity, rotation and position curves of EuRoc V1-01 easy dataset.
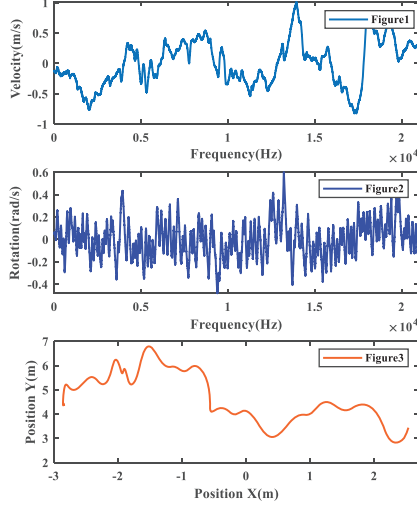
Fig. 2. Velocity, Rotation and Position curves of EuRoc dataset

Fig.3 demonstrates position curves and corresponding reinforcement learning feedback reward curves. Different motion sequences represent different rotation and translation incentives. Through the interaction of reinforcement learning, the reward brought by the excitation of each motion sequence is learned. Then select the best motion sequence for subsequent calibration. In this figure, the operating interval of the red dashed line can be considered as an optimal calibration sequence.
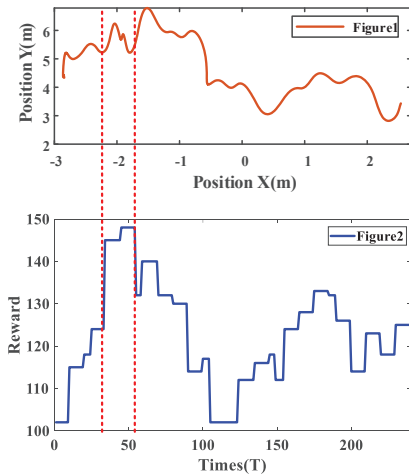
Fig. 3. Position curves and corresponding reinforcement learning feedback reward curves

Fig.4 and Fig. 5 demonstrates the translation error curves and rotation error curves respectively. In Fig. 4, the horizontal

axis represents time t_k, vertical axis is the rotation error in X, Y and Z axis. In Fig.5, the vertical axis is the translation error of roll, pitch, yaw. The blue line is the error curve and the red line is the 3σ bounds.
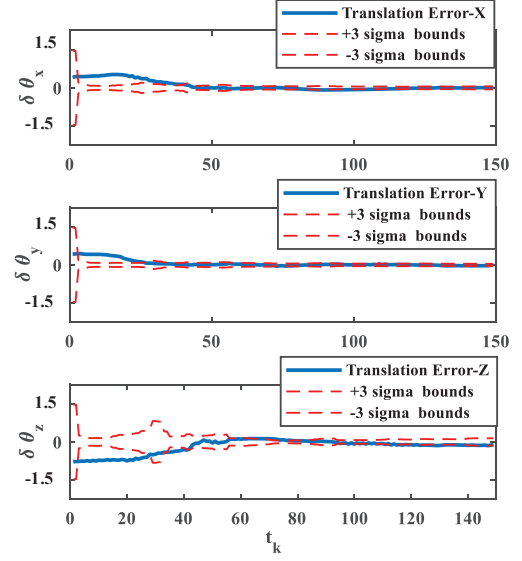
Fig. 4. Calibration translation error curves
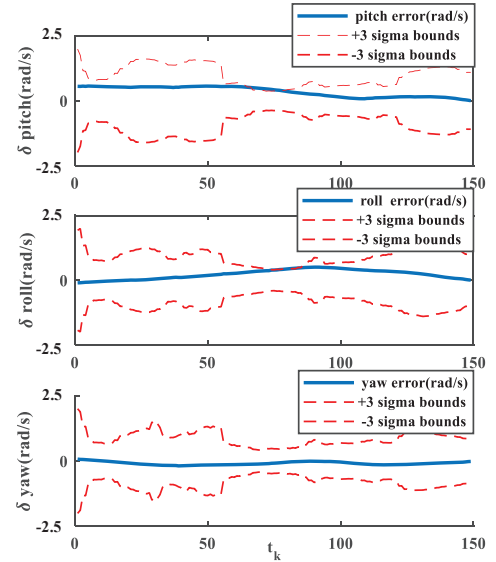
Fig. 5. Calibration rotation error curves

According to the analysis of the experimental results, the translation error of online relative pose calibration is 0.035m, and the rotation error is 0.399°. The proposed visual inertial calibration method based on reinforcement learning has reached the accuracy of the existing offline calibration algorithm and online self-calibration algorithm.

V. CONCLUSION

To serve the high-precision calibration parameters in VINS, a novel visual inertial self-calibration method based on reinforcement learning is proposed to settle the problems of complex calibration process and large amount of calculations. The major work and conclusions of this paper are drawn as follows:

1) To analyze the observability of the calibration parameters in the visual inertial system，FIM matrix is used to establish

the connection between matrix rank and nonlinear observability.

2) To reduce the amount of calculation in the calibration process, an automatic visual inertial self-calibration method based on reinforcement learning is proposed. This learning process trains the corresponding motion sequence of complete calibration parameter spatial regression and learns the optimal motion sequence that can regress the calibration parameter space.

3) The proposed method RL-VIC has a translation error of 0.035m and a rotation error of 0.399°. Compared with traditional calibration using all motion data, the calculation cost is greatly saved.

REFERENCES

[1] Cadena C, Carlone L, Carrillo H, et al. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age[J]. IEEE Transactions on Robotics, 2016, 32(6):1309-1332.

[2] Chen C, Zhu H, Wang L, Liu, Y. A Stereo Visual-Inertial SLAM Approach for Indoor Mobile Robots in Unknown Environments Without Occlusions[J]. IEEE Access, 2019: 185408-185421.

[3] Lin Y, Gao F, Qin T, et al. Autonomous aerial navigation using monocular visual-nertial fusion[J]. Journal of Field Robotics, 2018, 35(1): 23-51.

[4] Li C, Yu L, Fei S. Real-Time 3D Motion Tracking and Reconstruction System Using Camera and IMU Sensors[J]. IEEE Sensors Journal, 2019, 19(15): 6460-6466.

[5] Piao J, Kim S. Real-Time Visual–Inertial SLAM Based on Adaptive Keyframe Selection for Mobile AR Applications[J].IEEE Transactions on Multimedia, 2019, 21(11): 2827-2836.

[6] Kelly J, Sukhatme G S. Visual-Inertial Sensor Fusion: Localization, Mapping and Sensor-to-Sensor Self-calibration[J]. The International Journal of Robotics Research, 2011, 30(1): 56-79.

[7] Nobre F, Heckman C. Learning to calibrate: Reinforcement learning for guided calibration of visual-inertial rigs:[J]. The International Journal of Robotics Research, 2019: 1388-1402.

[8] Rehder J, Nikolic J, Schneider T, Hinzmann T, Siegwart R. Extending kalibr: Calibrating the extrinsics of multiple IMUs and of individual axes[C]. International Conference on Robotics and Automation, 2016: 4304-4311.

[9] Rehder J, Siegwart R. Camera/IMU Calibration Revisited[J]. IEEE Sensors Journal, 2017, 17(11): 3257-3268.

[10] Rehder J. Multi-sensor system calibrations [D]. Zurich: ETH Zurich, 2018.

[11] Li M, Mourikis A I. High-precision, consistent EKF-based visual-inertial odometry[J]. The International Journal of Robotics Research, 2013, 32(6): 690-711.

[12] Li M , Mourikis A I . Online temporal calibration for camera-IMU systems: Theory and algorithms[J]. International Journal of Robotics Research, 2014, 33(7):947-964.

[13] Kelly J, Sukhatme G S. Visual-inertial simultaneous localization, mapping and sensor-to-sensor self-calibration[C]. International Conference on Computational Intelligence in Robotics and Automation, 2009: 360-368.

[14] Mirzaei F M, Roumeliotis S I. A Kalman Filter-Based Algorithm for IMU-Camera Calibration: Observability Analysis and Performance Evaluation[J]. IEEE Transactions on Robotics, 2008, 24(5): 1143-1156.

[15] Tang H, Liu Y, Wang H. Constraint Gaussian Filter With Virtual Measurement for Online Camera-Odometry Calibration[J]. IEEE Transactions on Robotics, 2018:1-15.

[16] Yang Z, Shen S. Monocular Visual–Inertial State Estimation With Online Initialization and Camera–IMU Extrinsic Calibration[J]. IEEE Transactions on Automation Science and Engineering, 2017, 14(1): 39-51.

[17] Panahandeh G, Jansson M, Handel P. Calibration of an IMU-Camera Cluster Using Planar Mirror Reflection and Its Observability Analysis[J]. IEEE Transactions on Instrumentation and Measurement, 2015, 64(1): 75-88.

[18] Jauffret C. Observability and fisher information matrix in nonlinear regression[J]. Aerospace & Electronic Systems IEEE Transactions on, 2007, 43(2):756-75

Geomagnetic Matching Navigation for Aircraft Based on Improved Iterative Closest Contour Point Algorithm

Ninghui Xu
Key laboratory of micro-inertial instrument and advanced navigation technology,Ministry of education
School of Instrument Science and Engineering
Southeast University,
Nanjing, China
220193373@seu.edu.com

Lihui Wang *
Key laboratory of micro-inertial instrument and advanced navigation technology,Ministry of education
School of Instrument Science and Engineering
Southeast University
Nanjing, China
wlhseu@163.com

Abstract—Focusing on the problem of accumulated INS error of high-speed aircraft at high-altitude with unavailable satellite navigation, an improved iterative closest contour point (ICCP) algorithm is proposed for geomagnetic matching navigation helping eliminate INS errors. Aiming at local optimal problem of traditional ICCP with large initial positioning error, proposed algorithm improves its decision indicator based on multi-attribute decision-making. Matching effect is evaluated comprehensively by algorithm convergence degree indicator together with introduced trajectory correlation indicator. After judging the convergence degree of matching trajectory, its correlation with real trajectory is depicted by their geomagnetic intensity and linear correlation. Then, the fusion of two indicators based on entropy weight method is adopted as comprehensive decision indicator. Improved ICCP is compared with particle swarm optimization algorithm (PSO) to verify its matching effect in conditions of different map resolution and geomagnetic measurement noise. Experiment results indicate that improved algorithm's performance is significantly better than PSO and traditional ICCP with higher matching accuracy and better stability, helping optimize and decrease matching error of ICCP from thousands of meters to 100 meters.

Keywords—Geomagnetic navigation, geomagnetic matching, iterative closest contour point, multi-attributes decision making

I. INTRODUCTION

For high-speed aircraft at high-altitude, INS accumulated error can be improved by geomagnetic matching navigation when satellite navigation is unavailable. As an auxiliary navigation method, geomagnetic matching navigation can be integrated with INS, forming a totally autonomous navigation system which has been widely used [1,2]. In recent years, geomagnetic navigation technology has been applied and developed mainly in the field of underwater navigation [3-6], but less combined with navigation at high-altitude.

The geomagnetic intensity changes smoothly with space at high altitude [7], resulting in insufficient geomagnetic information. However, for high-speed carriers, a long trajectory can be obtained in a short period of time to ensure the richness of collected geomagnetic information, partly offsetting the bad matching suitability brought about by high-altitude. So geomagnetic matching navigation is feasible for high-speed aircraft at high-altitude.

As a commonly used geomagnetic matching algorithm, ICCP can eliminate INS positioning error and heading error at the same time [8]. Li et al. [9] deeply analyzed the impact of factors, including measurement noise of sensors, initial INS error, map resolution and matching path length, on the

performance of ICCP algorithm in matching navigation. Through simulations, the conclusion was achieved that ICCP matching error is heavily affected by measurement noise and initial INS error, and trajectories are easily mismatched if the map resolution is too high compared with the length of matching trajectories.

To solve the problem that geomagnetic measurement noise affects ICCP matching accuracy heavily and may lead to mismatch, Xiao et al. [10] generated several effective pseudo geomagnetic measurements within a confidence range according to statistical properties of measurement errors. Then ICCP was conducted on pseudo measurements to find possible positions of the carrier and matching results were limited by carrier's kinematical constraints to estimate its final position. The proposed algorithm improved the matching accuracy with measurement noise effectively. In addition, the geomagnetic measurement noise can also be suppressed by corresponding filtering [11]. Aiming at ICCP easily falling into local optimum with large initial INS positioning error, decreasing matching accuracy, Zhang et al. [12] used the position increment between two adjacent matching points to update the positioning result of geomagnetic/INS integrated navigation. And the effectiveness of the algorithm was verified by a semi-physical simulation experiment.

Multi-attribute decision-making has been used in matching navigation in recent years. Wang et al. [13] proposed a fuzzy multi-attribute decision making method to evaluate and select underwater terrain matching areas based on vague sets of multi topographic factor eigenvalues, overcoming the deficiency of fuzzy sets with single factor.

In general, ICCP geomagnetic matching algorithm requires the bias between INS trajectory and real trajectory to be small, so that the algorithm is easily falling into local optimum with a large INS initial error, leading to mismatch and low matching accuracy. Selecting the best trajectory from multi-match is usually used for reducing the impact of initial errors and improving matching performance. However traditional ICCP algorithm based on the decision indicator with single index cannot achieve optimal selection of matching trajectories especially when the resolution of the magnetic map is too high relative to the matching trajectory length due to the algorithm mechanism.

In order to solve the problem that traditional ICCP with decision indicator based on single index is greatly affected by initial error especially in high-resolution magnetic map, a comprehensive decision indicator based on multi-attribute

---

decision-making with both algorithm convergence degree and trajectory correlation taken into account is proposed. And geomagnetic matching navigation of high-speed aircraft in different geomagnetic map resolution with different geomagnetic measurement noise are simulated respectively to testify effectiveness and stability of the proposed algorithm. In addition, a new geomagnetic matching algorithm based on PSO is used and compared with proposed algorithm. Results show that the matching accuracy is effectively enhanced by improving the decision indicator of algorithm.

II. ICCP GEOMAGNETIC MATCHING ALGORITHM

Geomagnetic matching algorithm is essentially a digital map matching algorithm based on batch processing. Pre-matching trajectories are generated by transforming the INS indicated trajectory in search region according to geomagnetic intensity sequence based on trajectory transformation models. The search region is compressed with constraint conditions to improve the search efficiency and algorithm convergence speed. Then, an optimal trajectory is selected as output and INS error can be eliminated by trajectory transformation model parameters.

A. Algorithm Principle

ICCP rotates and translates INS indicated trajectory iteratively to approach the contours of corresponding geomagnetic intensity based on two-dimensional rigid transformation, so as to correct the cumulative error of INS. The principle is shown in Fig. 1.

For each match, the real trajectory is $X_i (i = 1,2, ... , N)$, where N is the number of trajectory points. INS indicated trajectory (P_i) is regarded as initial trajectory for matching, and its corresponding geomagnetic sequence is C_i. Generally, we suppose X_i to be on the contours of C_i, so we set the closest contour points sequence (Y_i) of X_i on C_i as our matching target. Then the trajectory is updated iteratively by rigid transformation with minimum Euclidean distance of Y_i and P_i criterion, until the termination condition is reached. The termination condition of iterations is related to its converge effectiveness, and the algorithm is normally assumed to converge locally if the trajectory almost keeps the same after transformation.

B. Search Closest Contour Points

Searching closest contour points is the prerequisite to obtain rigid transformation parameters. In order to decrease redundant matching and enhance algorithm efficiency, the search region of closest contour points is limited by initial INS positioning error constraint and kinematical constraint.
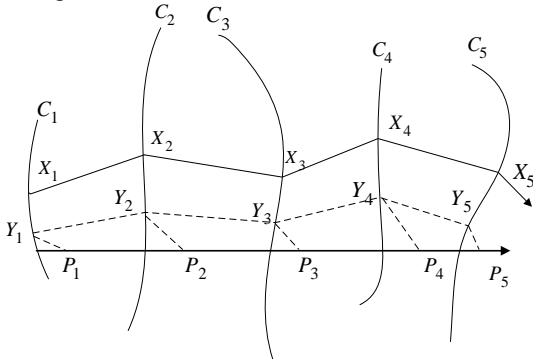
Fig. 1. Illustration of the principle of ICCP.
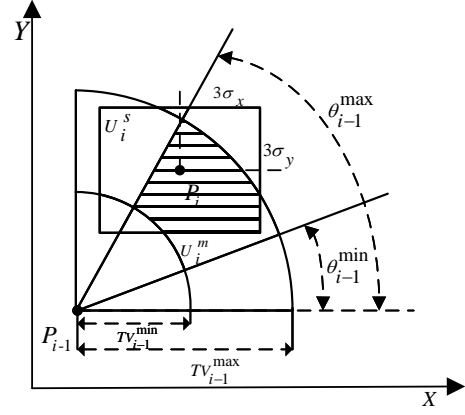
Fig. 2. Illustration of the search region for closest contour point.

Assuming that INS positioning error follows a standard normal distribution and its variance of east and north are σ_x^2 and σ_y^2 respectively. According to 3σ criterion, the closest contour point search region (U_i^s) of the trajectory point (P_i) can be constrained by INS positioning error:

$$U_i^s = \left\{(x,y) \big| |x - x_i| < 3\sigma_x, \ |y - y_i| < 3\sigma_y \right\} \qquad (1)$$

The search region can be re-constrained by carrier's kinematics of previous moment. With consideration of INS error and carrier performance limitation, the ranges of its speed and heading of previous moment are $[V_{i-1}^{min}, V_{i-1}^{max}]$ and $[\theta_{i-1}^{min}, \theta_{i-1}^{max}]$. The search region ($U_i^m$) constrained by kinematics is given by:

$$U_i^m = \Big\{(x,y) \big| (x,y) = (x_{i-1}, y_{i-1}) + TV_{i-1}(\cos(\theta_{i-1}), \sin(\theta_{i-1})),$$
$$V_{i-1} \in \left[V_{i-1}^{min}, V_{i-1}^{max}\right], \ \theta_{i-1} \in \left[\theta_{i-1}^{min}, \theta_{i-1}^{max}\right] \Big\} \qquad (2)$$

where T is the time interval.

So the search region (U_i) constrained by above two conditions is $U_i^s \cap U_i^m$, shown as shading part in Fig. 2.

Since geomagnetic map is a grid map, the search range of closest contour points is expanded iteratively in unit of minimum grid with trajectory points as center after search region is determined. As a result, the algorithm is easy to fall into local optimum when initial positioning error is large, especially in the geomagnetic map with relatively higher resolution. Such deficiency can be improved by selecting the best trajectory from multi-match results according to their decision indicator.

The pipeline of ICCP geomagnetic matching algorithm is shown in Fig. 3.

III. IMPROVED DECISION INDICATOR

ICCP's search strategy for closest contour points can lead to local optimum when ensuring the search accuracy, making the algorithm sensitive to initial positioning error so that the matching error can be extremely large when initial trajectory is far from real trajectory. However, traditional ICCP algorithm based on decision indicator donated singly by convergence exacerbates such effect, which need to be improved to enhance the matching accuracy.
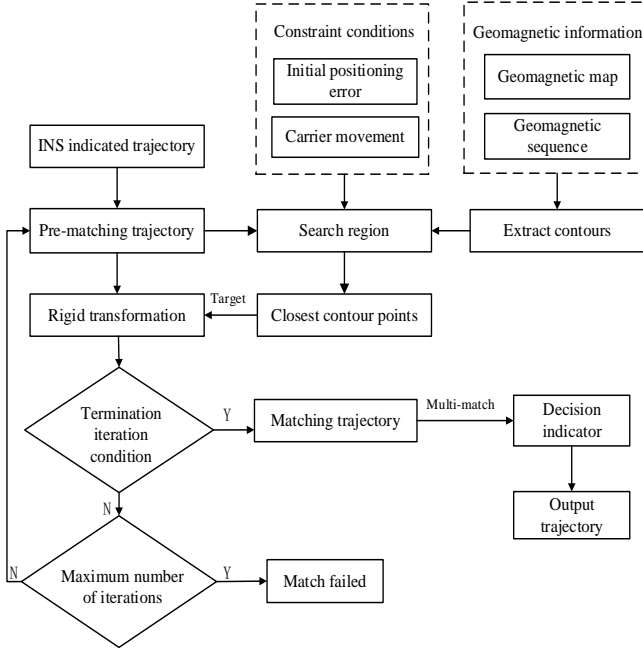
Fig. 3. A block diagram illustrating the pipeline of ICCP.

A. Traditional Decision Indicator

Traditional ICCP uses convergence degree of the algorithm as the indicator to evaluate its matching effect. INS trajectory is multi-matched to obtain M matching trajectories. The convergence degree indicator (D_j) of the j-th matching is depicted by average Euclidean distance of output trajectory (Y_j) between its closest contour points sequence (S_j). The convergence degree is higher with smaller D_j.

$$D_j = \frac{1}{N} \sum_{i=1}^{N} w_{ji} \| Y_{ji}(x) - S_{ji}(x) \| \qquad (3)$$

Where w_{ji} is the weight of the i-th trajectory point of the j-th match, calculated by distance between pre-matching trajectory (P_j) and its closest contour points sequence (S_j):

$$\begin{cases} \omega_{ji} = 1 - d_{ji} / d_{max} \\ d_{ji} = \| P_{ji} - Y_{ji} \| \end{cases} \qquad (4)$$

B. Improved Decision Indicator Based on Multi-attributes Decision-making

Decision indicator singly donated by convergence degree index cannot make effective evaluation for matching trajectories in high-resolution geomagnetic map with large initial error, because the algorithm can always achieve good convergence when the search unit is small. Therefore, with introduction of correlation between matching trajectories and real trajectory, an evaluation indicator based on multi-attribute decision-making is proposed in this paper to evaluate matching results comprehensively.

The trajectory correlation is depicted by two indexes from two aspects: geomagnetic intensity correlation and linear correlation.

Mean absolute difference (MAD) is selected as geomagnetic intensity correlation criterion. The intensity correlation index (I_j) is calculated by MAD of geomagnetic measurement sequence (B_R^{ji}) and geomagnetic sequence

(B_S^{ji}) of matching trajectory. Bilinear interpolation is used here for matching trajectory to obtain its geomagnetic sequence. The intensity correlation between matching trajectory and real trajectory is better with smaller I_j.

$$I_j = \frac{1}{N} \sum_{i=1}^{N} \left| B_R^{ji} - B_S^{ji} \right| \qquad (5)$$

Linear correlation actually reveals the correlation of geomagnetic intensity fluctuation on trajectories. The linear correlation index (R_j) is calculated by:

$$R_j = \frac{\sqrt{Var(B_R^j) \cdot Var(B_S^j)}}{Cov(B_R^j, B_S^j)} \qquad (6)$$

Trajectories with smaller R_j has better linear correlation with real trajectory. After being normalized to unified scale, I_j and R_j constitute trajectory correlation indicator (C_j). The normalization and C_j is given by:

$$\begin{cases} I'_j = \dfrac{I_j - I_{min}}{I_{max} - I_{min}} \\ \\ R'_j = \dfrac{R_j - R_{min}}{R_{max} - R_{min}} \end{cases} \qquad (7)$$

$$C_j = I'_j + R'_j \qquad (8)$$

Then comprehensive evaluation indicator (F) can be obtained by fuse convergence degree indicator (D_j) with trajectory correlation indicator (C_j) by entropy weight method. Entropy weight method determines the weight of each indicator according to its donated information amount, which is measured based on the data differences in M matching trajectories from multi-match.

Indicators also need to be normalized to unified scale in the same way before being weighted. Taking convergence degree indicator (D) as an example, after being normalized to D', its information entropy (e_D) and weight (w_D) are:

$$e_D = -\frac{1}{\ln(2)} \sum_{j=1}^{M} D'_j \ln(D'_j) \qquad (9)$$

$$w_D = \frac{1 - e_D}{2 - (e_D + e_C)} \qquad (10)$$

Where e_C is information entropy of trajectory correlation indicator (C). Then comprehensive indicator (F_j) of the j-th matching trajectory is given by:

$$F_j = w_D \cdot D'_j + w_C \cdot C'_j \qquad (11)$$

C. Comparative Verification

TABLE I.　MATCHING RESULTS

Map resolution (°)	Matching error (m)	
	ICCP	Improved ICCP
0.05	406.09	283.79
0.005	3521.05	105.99
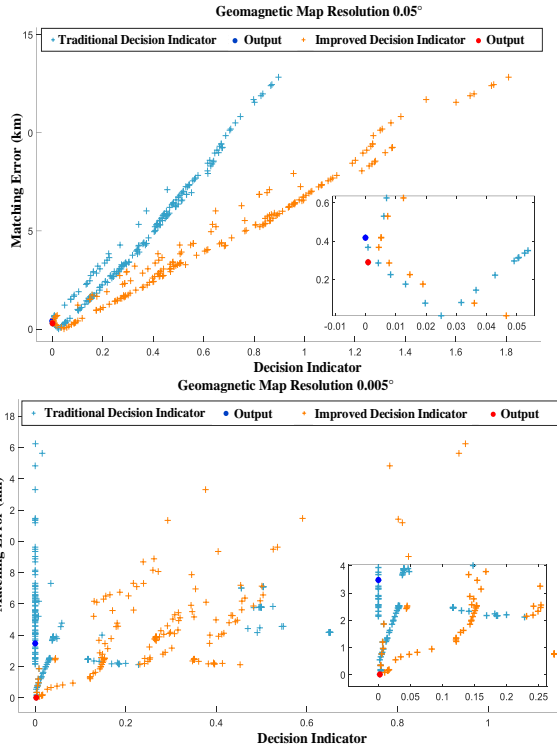
Fig. 4. Distribution of matching error and decision indicators.

In order to verify the impact of decision-making indicators on matching accuracy, traditional ICCP and improved ICCP are used to conduct geomagnetic matching navigation in geomagnetic maps of 0.05 ° and 0.005 ° resolution. Matching results are shown in TABLE. I.

Each map is multi-matched for 200 times. Decision indicators value and matching error of each matching trajectories are recorded. Their distribution is shown in Fig.4.

The distribution of traditional and improved decision indicators with matching error in map with resolution of 0.05°illustrates that, the correlation between two decision indicators and matching error is obviously good. And the local enlarged image indicates that comprehensive decision indicator can help select the optimal trajectory.

As map resolution increased to 0.005°, the distribution illustrates that, due to the limitation of ICCP's search strategy for closest contour points, most matching converge well, making convergence singly donated traditional indicator cannot achieve comprehensive evaluation of trajectories. While the correlation between multi-attribute indicator and matching error is still good. The matching results in TABLE I reveal that, with both convergence and trajectory correlation taken into account, multi-attribute decision-making based improved ICCP can improve matching accuracy effectively. Matching error decreases from 3521.05m to 105.99m.

Parameter	Value	Unit
Speed	4	Mach
Initial position error	[3000,5000]	m
Gyro rift	0.01	°/h
Accelerometer bias	50	ug
Sampling period	2	s

IV. SIMULATION AND ANALYSIS

Particle swarm optimization (PSO) geomagnetic matching algorithm is a new geomagnetic matching method [14-16], which corrects INS errors by optimizing the parameters of trajectory transformation model. Compared with incremental strategy of ICCP, PSO uses global search strategy to search for target trajectories in the whole region under constraint conditions, so it is less affected by initial INS positioning error, and the matching accuracy is generally better than traditional ICCP. In order to verify the matching effect of improved ICCP geomagnetic matching algorithm, it is compared with both traditional ICCP and PSO matching algorithm, of which the particle number was set to 1000 to improve its matching accuracy, so as to achieve an effective verification.

Enhanced Magnetic Model (EMM) has high spatial resolution. EMM2017 was taken as the geomagnetic map, ranging from 38°to 40°in latitude, from 120°to 122°in longitude and at 40 km in height.

Traditional ICCP, improved ICCP and PSO are used to conduct geomagnetic matching navigation of high-speed aircraft at high altitude respectively in conditions of different map resolution and different geomagnetic measurement noise. And each condition is simulated for 100 times with 200 multi-match. When matching error is less than 1000m, the matching is supposed to be success. The parameters of the simulation are shown in TABLE. II, and trajectory points number is 15.

A. Geomagnetic Matching in Different Resolution

In order to verify the matching effect of improved ICCP in different resolution, geomagnetic maps with resolution of 0.05°, 0.01°, 0.005°, 0.00125°are selected for simulation. Matching results are shown in TABLE. III, from which we can see that traditional ICCP has an exceptional matching performance in 0.05° resolution map, while it is completely ineffective when the resolution rise to 0.005° and 0.00125°. It's apparent that PSO and improved ICCP have better matching accuracy than traditional ICCP, so we mainly analyze the proposed algorithm by comparing it with PSO. Fig.5. shows the matching results of PSO and improved ICCP.

According to the comparison of two algorithms, improved ICCP has a better performance in matching than PSO obviously. As map resolution increases from 0.005° to

TABLE II.　　SIMULATION PARAMETERS

TABLE III.　　MATCHING RESULTS IN DIFFERENT RESOLUTION

Map resolution (°)	Success rate (%)			Matching error (m)		
	ICCP	PSO	Improved ICCP	ICCP	PSO	Improved ICCP
0.05	100	56.0	100.0	98.99	332.29	51.75
0.01	58	72.0	100.0	466.37	326.80	65.00
0.005	26	99.0	100.0	876.35	217.27	118.10
0.00125	3	64.0	87.0	897.58	360.49	345.72

TABLE IV.   MATCHING RESULTS WITH DIFFERENT NOISE

Standard deviation of the noise (nT)	Success rate (%)			Matching error (m)		
	ICCP	PSO	Improved ICCP	ICCP	PSO	Improved ICCP
0	26	99.0	100.0	876.35	217.27	118.10
1	5	95.0	99.0	835.67	232.26	152.51
2	0	99.0	93.0	—	285.48	170.29
3	0	95.0	96.0	—	355.11	264.40
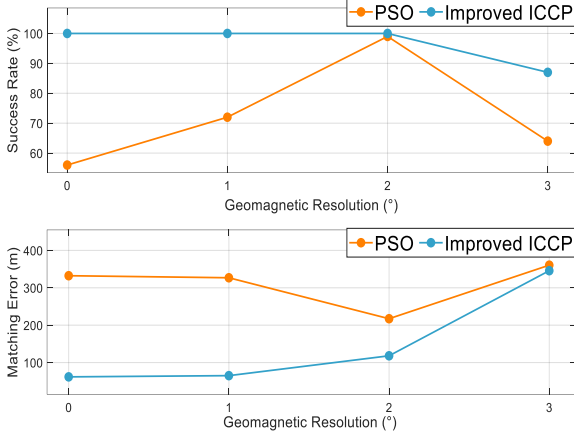
Fig. 5.   Matching results in different resolution of PSO and improved ICCP

0.00125°, initial positioning error becomes larger relative to grid distance, resulting in matching error of improved ICCP increasing from 51.75m to 345.72m gradually. And matching error of PSO is stable (about 300m), whereas its matching success rate is lower than improved ICCP. Especially when map resolution is 0.005°, the matching success rate of improved ICCP reaches 100% compared with 56% of PSO. In general, improved ICCP geomagnetic matching algorithm is effective with different geomagnetic map resolution.

B. Geomagnetic Matching with Different Measurement Noise

In order to verify the stability of improved ICCP, considering that geomagnetic intensity changes smoothly with space and geomagnetic measurement noise is small at high altitude, measurement noise with mean value of 0nt and standard deviation of 0-3nt are added to geomagnetic sequence. The simulation is conducted in geomagnetic map with resolution of 0.005° and TABLE. IV shows the results of three matching methods. In view of that traditional ICCP is ineffective with any noise, we still analyze improved ICCP and PSO based on their matching results. Fig. 6. shows the matching results of PSO and improved ICCP.

It can be seen that with different noise in geomagnetic measurement, the matching error of improved ICCP is still smaller than PSO, and their gap is stable round 100m. As noise increases from 0nt to 3nT, the matching accuracy of improved ICCP decreases slightly with matching error increasing from 118.1m to 264.40m, indicating that improved ICCP geomagnetic matching algorithm is steady with external interference.

V.   CONCLUSION

With background of geomagnetic matching navigation for high-speed carrier at high altitude, this paper analyzes the impact of initial positioning error and geomagnetic map resolution on matching accuracy from algorithm's search
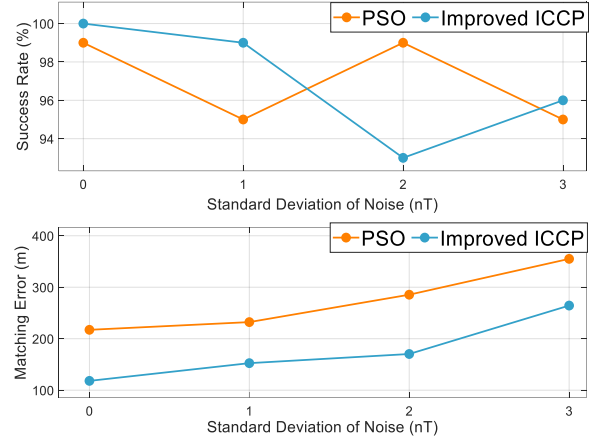
Fig. 6.   Matching results with different noise of PSO and improved ICCP

strategy for closest contour points and trajectory decision indicator. Trajectory correlation is introduced together with convergence indicator of traditional algorithm to evaluate and select matching trajectory for output from multi-match. And correlation indicator is composed of geomagnetic intensity correlation index and linear correlation index. Experimental results show that, based on improved decision indicator, the output trajectory selected from multi-match is closer to real trajectory. Simulations in different conditions indicate that the matching accuracy and stability of improved ICCP are better than that of PSO based on global search strategy. So the improved algorithm can eliminate the accumulated INS error of high-speed aircraft effectively and steadily.

REFERENCES

[1]   S. Yeh, W. Hsu, W. Lin and Y. Wu, "Study on an Indoor Positioning System Using Earth's Magnetic Field," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 3, pp. 865-872.

[2]   G. Huang, B. K. Taylor and D. Akopian, "A Low-Cost Approach of Magnetic Field-Based Location Validation for Global Navigation Satellite Systems," in IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 12, pp. 4937-4944.

[3]   H. Li, M. Y. Liu and K. Liu, "Bio-inspired geomagnetic navigation method for autonomous underwater vehicle," Journal of Systems Engineering and Electronics, 2017, vol. 06, pp. 1203-1209.

[4]   C. V. Kaji, R. C. Hoover and S. Ragi, "Underwater Navigation using Geomagnetic Field Variations," 2019 IEEE International Conference on Electro Information Technology (EIT), Brookings, SD, USA, 2019, pp. 1-6.

[5] L. H. Wang, L. Yu, N. Qiao, D. S. Sun, "Analysis and Simulation of Geomagnetic Map Suitability Based on Vague Set," 2016, vol. 5, pp. 1114-1124.

[6] J. Sun, C. H. Song, X. H. Xiao, X. M. Jin, J. H. Ni, "Calculation Error's Correction of Three-Component Geomagnetic Field's Marine Survey," Advanced Materials Research, 2013.

[7] T. Zha, X. N. zhu, D. F. Cheng, Z. J. Zhou, "Geomagnetic field spatial difference measuring method for underwater geomagnetic navigation technology," Journal of Jilin University, 2017, vol. 1, pp. 316-322.

[8] Y. Lin, L. Yan, Q. X. Tong, "Underwater geomagnetic navigation based on ICP algorithm," 2007 IEEE International Conference on Robotics and Biomimetics (ROBIO), Sanya, 2007, pp. 2115-2120.

[9] M. Li, Y. Liu and L. Xiao, "Performance of the ICCP algorithm for underwater navigation," 2014 International Conference on Mechatronics and Control (ICMC), Jinzhou, 2014, pp. 361-364.

[10] J. Xiao, X. S. Duan, X. H. Qi, "Iterated closest contour point algorithm for geomagnetic matching based on probability data association," in Journal of Chinese Inertial Technology, 2018, vol. 2, pp. 202-208.

[11] N. Qiao, L. H. Wang, Q. Y. Liu, H. Q. Zhai, "Multi-scale eigenvalues Empirical Mode Decomposition for geomagnetic signal filtering," in Measurement, 2019, vol. 11, pp. 885-891.

[12] Y. Zhang, C. Li, S. Liu, M. Jiang and L. Guo, "Locating method of geomagnetic/inertial integrated navigation system by forecasting the geomagnetic matching initial value," Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference, Yantai, 2014, pp. 235-239.

[13] L. H. Wang, D. S. Sun, Q. Y. Liu, L. Yu, "Matching area selection of an underwater terrain navigation database with fuzzy multi-attribute decision making method," Proceedings of the Institution of Mechanical Engineers Part M-Journal of Engineering for the Maritime Environment, 2019, vol. 4 pp. 1130-1140.

[14] Z. Wu, Y. Wu, X. Hu and M. Wu, "Calibration of Three-Axis Magnetometer Using Stretching Particle Swarm Optimization Algorithm," in IEEE Transactions on Instrumentation and Measurement, vol. 62, no. 2, pp. 281-292.

[15] N. Qiao, L. H. Wang, D. S. Sun, M. Z. Ma, L. Yu, "Application of particle swarm optimization in track planning," in Journal of Chinese Inertial Technology, 2018, vol. 06, pp. 787-791.

[16] L. Zhou, X. H. Cheng, "Seabed terrain-aided SINS location based on constrained particle swarm optimization method," in Journal of Chinese Inertial Technology, 2015, vol. 03, pp. 369-372+384.

An Improved Interacting Multiple Model Algorithm for INS/DVL Integrated Navigation System

Lanhua Hou
School of Instrument Science and Engineering
Southeast University
Nanjing, China
lanhua_hou@163.com

Xiaosu Xu*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
xxs@seu.edu.cn

Yiqing Yao
School of Instrument Science and Engineering
Southeast University
Nanjing, China
yucia@sina.com

Di Wang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
wangdigood2011@163.com

Abstract—**In Inertial Navigation System (INS)/ Doppler Velocity Log （DVL） integrated navigation system, due to the time-varying measurement noise，the inaccurate measurement model results in substantial estimation errors or even filter divergence. To address this problem, Interacting Multiple Model (IMM) algorithm which uses more than one model is always considered. The efficiency of IMM algorithm depends on the accuracy of model set. In order to cover the real model with smaller model set, an Improved IMM(IIMM) algorithm is proposed in this paper. A variable model set based on the model probability weighted average of the model parameter is generated to pursue the real model. Meanwhile, the time-varying state transition probability matrix based on model probability difference is introduced to accelerate the switch of model set to the true model. Simulation and experiment indicate that the model set can cover the true model rapidly and precisely with only three models, and the accuracy of the SINS/DVL integrated navigation system is improved by applying the proposed method.**

Keywords—Interacting Multiple Model, integrated navigation, variable model set, state transition probability matrix

I. INTRODUCTION

According to statistics, ocean where is rich in mineral resources, takes up three-fourths of the earth's surface. Underwater navigation technology, as the core technology of ocean exploration, is the most difficult to tackle. In the domain of underwater navigation, Inertial Navigation System (INS) with the performance of high autonomy and anti-interference is widely used. But there are some performance deficiencies for INS (e.g. unbounded position error growth and position/velocity/attitude Schuler Oscillations). In light of this, acoustic positioning system, Doppler Velocity Log (DVL), Global Position System(GPS), depthmeter, and terrain matching are commonly introduced as alternative to provide navigation information for underwater vehicles[1, 2]. Among these, DVL has become one of the main auxiliary navigation devices in underwater navigation by virtue of its reliability, autonomy and convenience. Recently, INS/DVL integrated system has developed into a reliable and important navigation system in underwater navigation[3].

In INS/DVL integrated system, Kalman filter(KF) is usually employed to blend navigation data from INS and DVL[4]. The optimality of KF relays on the correct prior

knowledge of the process noise covariance matrix Q and the measurements noise covariance R[5]. The matrix Q is often regarded as constant because the process noise in the navigation process is approximately invariable. Nevertheless, due to the complex underwater environment, such as: ocean currents, marine organism, changeable template and salinity, the random error of DVL is variable and unpredicted. The affiliated matrix R is not in compliance with reality and results in substantial estimation errors or even filter divergence. The problem of optimal filtering in the presence of unknown measurement noise has therefore received a lot of attention. There are two mainstream methods, Adaptive Kalman Filter (AKF) and Interacting Multiple Model(IMM) method.

The AKF algorithm estimates the statistical characteristics of process and measurement noises to adapt to the uncertain model[6]. The AKF can be divided into four categories: Bayesian, Maximum Likelihood (ML), correlation, and covariance matching. Based on these theories, AFK has been widely utilized in the integrated navigation systems[7, 8]. The IMM method gets around the difficulty due to the model uncertainty by using more than one model[9]. It was generated to handle the model uncertainty in the target tracking area. Recently, it has become popular in integrated navigation systems[10-12]. In this paper, the IMM method is mainly discussed.

To the best of our knowledge, the efficiency of IMM method depends on the accuracy of model set[13]. In order to cover the real model as much as possible, a bigger model set is needed[14,15]. Paradoxically, with the increase of the model set, the computation complexity increases, and the competition between sub models leads to the filter deterioration. To address this, Sage-Husa AKF was introduced to IMM to establish model set to adaptively cover the true model. And a switching criterion based Hybrid Interacting Multiple Model (HIMM) algorithm was also proposed[16]. However, additional calculation of Sage-Husa AKF and switching criterion in the two methods partly complicates this problem. To simplify the calculation and improve the estimation accuracy, an Improved Interacting Multiple Model(IIMM) algorithm is proposed in this paper. The model-probability-weighted-average-based adaptive model set is proposed to cover the real model. Constructed by the model probability weighted average of model parameter in the last epoch, the new model gradually approaches to the true model. Meanwhile, an adaptive transition probability matrix

Xiaosu Xu is the corresponding author. (e-mail: xxs@seu.edu.cn).

is introduced to accelerate the switch to the true model set when the true model is not in the current model set. The IIMM algorithm enables INS/DVL integrated navigation system to operate in uncertain, complex, and changing situations.

The structure of this paper is as follows. In section 2, an INS/DVL integrated system is designed. In section 3, the IIMM algorithm is specifically illustrated. In section 4, simulation and vehicle test are comprehensively conducted to illustrate the superiority of the proposed IIMM algorithm. Section 5 is devoted to conclusion.

II. INS/DVL INTEGRATED SYSTEM.

In INS/DVL integrated system, the loosely coupled method is frequently utilized because the information provided by DVL is the velocity of the vehicle rather than the original information. The diagram of the INS/DVL integrated system is shown in Fig. 1. Kalman filter is used to fuse the updated navigation error of INS and the velocity of DVL and then figures out the navigation error to correct the navigation information. To cast the integrated process in a Kalman filter framework, state equation and measurement equation are established by the time rate differential equations of INS and the velocity difference of INS and DVL in the body frame.
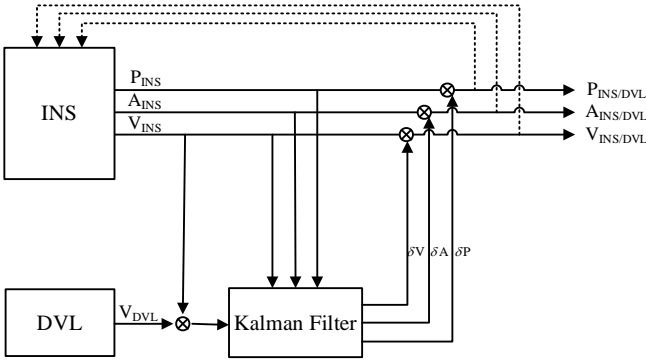
Fig. 1. Diagram of the INS/DVL integrated system.

A. State Equation

The state equation is described as follows:

$$\dot{\mathbf{X}} = \mathbf{F}\mathbf{X} + \mathbf{G}\mathbf{W} \tag{1}$$

where $\mathbf{X}$ is the state vector, $\mathbf{F}$ is the state transition matrix, $\mathbf{G}$ is the system noise matrix, and $\mathbf{W}$ is the process noise vector. The state vector $\mathbf{X}$ is defined as:

$$\mathbf{X} = \begin{bmatrix} \phi_x & \phi_y & \phi_z & \delta V_E^n & \delta V_N^n & \delta V_U^n & \delta\lambda & \delta L & \delta h \\ & \varepsilon_x & \varepsilon_y & \varepsilon_z & \nabla_x & \nabla_y & \nabla_z & K_d & \end{bmatrix}^T \tag{2}$$

where $\phi_{x,y,z}$ are misalignment angles, and $\delta V_{E,N,U}^n$ are east, north and upward velocity errors, respectively. $\delta\lambda$, δL, δh are the latitude, longitude, and height errors, respectively. $\varepsilon_{x,y,z}$ and $\nabla_{x,y,z}$ represent the accelerometer biases and gyro biases in three directions of the body frame. K_d is the DVL scale factor. The system state transition matrix $\mathbf{F}$, the matrix $\mathbf{G}$ and $\mathbf{W}$ can be expressed as same as [17].

B. Measurement Equation

The measurement equation is defined as follows[18]:

$$\mathbf{Z} = \hat{\mathbf{V}}_b - \mathbf{V}_{bDVL} = \mathbf{H}\mathbf{X} + \mathbf{V} \tag{3}$$

where $\mathbf{Z}$ is the measured matrix, $\hat{\mathbf{V}}_b$ is the velocity calculated by INS under frame b, $\mathbf{V}_{bDVL}$ is the velocity of DVL, $\mathbf{H}$ is the measurement transfer matrix, and $\mathbf{V}$ is the measurement information noise. The measurement transfer matrix $\mathbf{H}$ is deduced as:

$$\mathbf{H} = \begin{bmatrix} C_{31}v_N - C_{21}v_U & C_{11}v_U - C_{31}v_E & C_{21}v_E - C_{11}v_N & & -v_{xDVL} \\ C_{32}v_N - C_{22}v_U & C_{12}v_U - C_{32}v_E & C_{22}v_E - C_{12}v_N & C_n^b & \mathbf{0}_{3\times9} & -v_{yDVL} \\ C_{33}v_N - C_{23}v_U & C_{13}v_U - C_{33}v_E & C_{23}v_E - C_{13}v_N & & -v_{zDVL} \end{bmatrix} \tag{4}$$

III. THE IMPROVED INTERACTING MULTIPLE MODEL ALGORITHM

A. Interacting Multiple Model Algorithm

1) Interacting

The predicted model probability can be calculated as[19]:

$$\mu_{i \to j}(k-1) = p_{i \to j}\mu_i(k-1) / \sum_{i=1}^{3} p_{i \to j}\mu_i(k-1) \tag{5}$$

where $i, j = 1, 2, 3$. $p_{i \to j}$ is the transition probability between model m_i and model m_j from the last epoch $k-1$ to current epoch k. $\mu_i(k-1)$ denotes the model probability estimated at the end of the epoch $k-1$.

The mixing state vector $\hat{\mathbf{X}}_{Oj(k-1)}$ and its estimate covariance $\mathbf{P}_{Oj(k-1)}$ of filter j ($j = 1, 2, 3$) can be calculated:

$$\hat{\mathbf{X}}_{Oj(k-1)} = \sum_{i=1}^{3} \hat{\mathbf{X}}_{i(k-1)}\mu_{i \to j}(k-1) \tag{6}$$

$$\mathbf{P}_{Oj(k-1)} = \sum_{i=1}^{3} \left(\mathbf{P}_{i(k-1)} + \varpi\varpi^T \right)\mu_{i \to j}(k-1) \tag{7}$$

$$\varpi = \hat{\mathbf{X}}_{i(k-1)} - \hat{\mathbf{X}}_{Oj(k-1)} \tag{8}$$

where $\hat{\mathbf{X}}_{i(k-1)}$ and $\mathbf{P}_{i(k-1)}$ are the state estimate and its covariance matrix of filter i at the last epoch.

2) Model filtering

With the output of interacting process, conventional Kalman filtering is performed:

$$\hat{\mathbf{X}}_{j(k,k-1)} = \mathbf{\Phi}_{j(k,k-1)}\hat{\mathbf{X}}_{Oj(k-1)} \tag{9}$$

$$\mathbf{P}_{j(k,k-1)} = \mathbf{\Phi}_{j(k,k-1)}\mathbf{P}_{Oj(k-1)}\mathbf{\Phi}_{j(k,k-1)}^T + \mathbf{\Gamma}_{j(k)}\mathbf{Q}_{j(k)}\mathbf{\Gamma}_{j(k)}^T \tag{10}$$

$$\mathbf{K}_{j(k)} = \mathbf{P}_{j(k,k-1)}\mathbf{H}_{j(k)}^{T}\mathbf{A}_{j(k)}^{-1} \qquad (11)$$

$$\hat{\mathbf{X}}_{j(k)} = \hat{\mathbf{X}}_{j(k,k-1)} + \mathbf{K}_{j(k)}\mathbf{r}_{j(k)} \qquad (12)$$

$$\mathbf{P}_{j(k)} = \left(\boldsymbol{I} - \mathbf{K}_{j(k)}\mathbf{H}_{j(k)}\right)\mathbf{P}_{j(k,k-1)} \qquad (13)$$

$$\mathbf{r}_{j(k)} = \mathbf{Z}_{j(k)} - \mathbf{H}_{j(k)}\hat{\mathbf{X}}_{j(k,k-1)} \qquad (14)$$

$$\mathbf{A}_{j(k)} = \mathbf{H}_{j(k)}\mathbf{P}_{j(k,k-1)}\mathbf{H}_{j(k)}^{T} + \mathbf{R}_{j(k)} \qquad (15)$$

3) Model probability update

The likelihood function of each model can be derived as:

$$f_{j}\left(k\right) = \frac{\exp(-\frac{1}{2}\mathbf{r}_{j(k)}^{T}\mathbf{A}_{j(k)}^{-1}\mathbf{r}_{j(k)})}{\left(\left(2\pi\right)^{m}\left|\mathbf{A}_{j(k)}\right|\right)^{1/2}} \qquad (16)$$

where *m* is the dimension of the observation vector. Then, the model probability can be calculated as:

$$\mu_{j}\left(k\right) = \frac{f_{j}\left(k\right)\sum\limits_{i=1}^{3}p_{i \to j}\mu_{i}\left(k-1\right)}{\sum\limits_{j=1}^{3}f_{j}\left(k\right)\sum\limits_{i=1}^{3}p_{i \to j}\mu_{i}\left(k-1\right)} \qquad (17)$$

4) Estimation fusion

With the model probability, the state vectors and its covariance matrices can be achieved:

$$\hat{\mathbf{X}}_{k} = \sum_{j=1}^{3}\hat{\mathbf{X}}_{j(k)}\mu_{j}\left(k\right) \qquad (18)$$

$$\mathbf{P}_{k} = \sum_{j=1}^{3}\mu_{j}\left(k\right)\left(\mathbf{P}_{j(k)} + \left(\hat{\mathbf{X}}_{j(k)} - \hat{\mathbf{X}}_{k}\right)\left(\hat{\mathbf{X}}_{j(k)} - \hat{\mathbf{X}}_{k}\right)^{T}\right) \qquad (19)$$

B. Improved Interacting Multiple Model Algorithm

The IIMM algorithm with the adaptive model set construction and the adaptive transition probability matrix is shown in Fig .2.
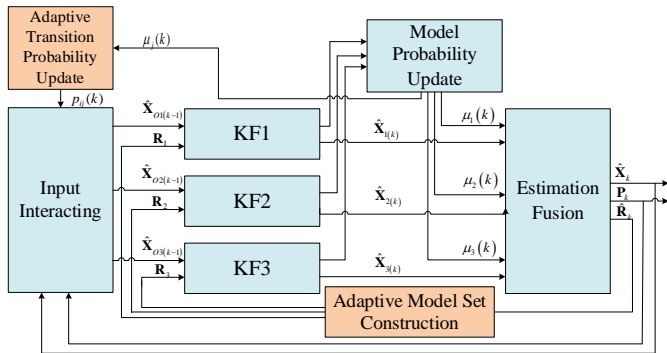
Fig. 2. Diagram of the IIMM algorithm.

1) Adaptive model set construction

The efficiency of IMM algorithm depends on the accuracy of model set. When the models in the model set match the true model appropriately, IMM performs well. However, once the true model is outside the model set, the accuracy of IMM cannot be guaranteed. Therefore, a model-probability weighted-average-based adaptive model set is proposed to establish new model set to cover true model as much as possible.

In IMM algorithm, the model probability represents the probability that the model is the true model. Accordingly, calculating the model probability weighted average of model parameter $\hat{\mathbf{R}}_{k}$:

$$\hat{\mathbf{R}}_{k} = \sum_{j=1}^{3}\mathbf{R}_{j(k)}\mu_{j}\left(k\right) \qquad (20)$$

Analytically, $\hat{\mathbf{R}}_{k}$ approximates the true measurement noise covariance matrix as much as possible.

Suppose there are three models $\mathbf{R}_{1}$, $\mathbf{R}_{2}$, $\mathbf{R}_{3}$.

$$\mathbf{R}_{1} = 1/\eta\hat{\mathbf{R}}_{k}, \quad \mathbf{R}_{2} = \hat{\mathbf{R}}_{k}, \quad \mathbf{R}_{3} = \eta\hat{\mathbf{R}}_{k} \qquad (21)$$

where η is constant and is set as 2 in this paper. $\mathbf{R}_{1}$, $\mathbf{R}_{2}$ and $\mathbf{R}_{3}$ are the small model, middle model and big model, respectively. When the true model is not covered in the model set, the calculated $\hat{\mathbf{R}}_{k}$ is closer to true value $\mathbf{R}_{k}$ than middle model $\mathbf{R}_{2}$.Then pick $\hat{\mathbf{R}}_{k}$ as the middle model and establish the mode set. We can see that the model set is closer to the true model. Repeat in this way until the model set covers the true model. Therefore, adaptive model switch is accomplished.

The switching process when the model set does not cover the real model is shown in Fig. 3. In Fig.3(a), when the real model is smaller than the model set, the probability weighted average value $\hat{\mathbf{R}}_{k}$ is smaller than the medium model in the last epoch. Pick $\hat{\mathbf{R}}_{k}$ as the middle model and establish the model set. Owing to the probability weighted average value, the middle model in current epoch is smaller than the last epoch. Accordingly, the model set becomes smaller and the adaptive switch of model set is implemented successively. It is similar when the real model is bigger than the model set and it is illustrated in Fig. 3(b).
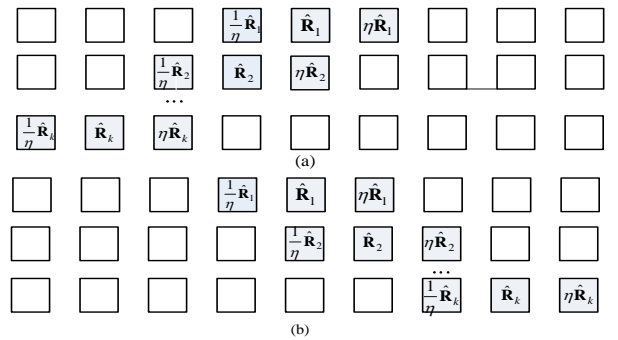
Fig. 3. (a)Diagram of the switching process when the real model is smaller than the model set. (b)Diagram of the switching process when the real model is bigger than the model set.

2) Adaptive transition probability matrix

In the conventional IMM algorithm, model transformation probability is fixed and determined according to experience. However, due to the variability and uncertainty of the real model, fixed model transformation probability is not in compliance with reality, which will lead to the Kalman filter deterioration. To address this, time-varying state transition matrix was pioneered for target tracking[20-22]. In the INS/DVL system, when the real model is not covered in the model set, the fixed model transformation probability may lead the model switching sluggish and then results in inaccurate filtering. In order to switch to the true model rapidly, the probability of the models which is far from the true model transforming to the model which is close to the real model should be larger. Therefore, an adaptive model transformation probability is introduced to IIMM algorithm for INS/DVL integrated navigation system in this paper.

The model transformation in IMM algorithm follows the Markov process, the transition probability between model m_i and model m_j from the last epoch $k-1$ to current epoch k is indicated as $p_{i \to j}$. The model probability $\mu_j(k)$ represents the matching degree between the model j and the real model. The greater the probability, the better the model j matches the real model, and the greater the probability that other models transfer to model j. Therefore, the model probability difference at adjacent moments can be used to correct the transition probability between the models. Calculating the probability difference at adjacent moments:

$$\Delta\mu_j(k) = \mu_j(k) - \mu_j(k-1), i,j = 1,2,3 \qquad (22)$$

To ensure positive definiteness of the revised transition probability, use the exponential form to correct transition probability:

$$p_{i \to j}(k) = \exp[\Delta\mu_j(k)]p_{i \to j}(k) \qquad (23)$$

It can be seen that, when the model probability of model j increases, $p_{i \to j}(k)$ increases, and the model switch accelerates, vice versa.

$$\begin{cases} \mu_j(k) - \mu_j(k-1) > 0, \ \Delta\mu(k) > 1, \ p_{i \to j}(k) \uparrow \\ \mu_j(k) - \mu_j(k-1) < 0, \ \Delta\mu(k) < 1, \ p_{i \to j}(k) \downarrow \end{cases} \qquad (24)$$

In order to ensure the sum of state transition probabilities to be one, it is needed to normalize the transition probability:

$$p_{ij}(k) = \frac{p_{ij}(k)}{\sum_{j=1}^{r} p_{ij}(k)} \qquad (25)$$

IV. SIMULATION AND VEHICLE TEST

The performance and comparison of the INS/DVL integrated system based on KF, Interacting Multiple Model Kalman Filter(IMMKF) and Improved Interacting Multiple Model Kalman Filter(IIMMKF) are evaluated on simulation and vehicle tests. The result is of great interest.

A. Simulation

The trajectory with straight line and two corners is simulated and the INS/DVL integrated system is applied to navigate. The simulated trajectory is shown in Fig .4. The start point is set as latitude 34° N and longitude 108° E. The initial attitude error is set as [0.1°, 0.1°, 0.3°]. The drift bias and the random walk noise of the accelerometer are set as $50\mu g$ and $50\mu g/\sqrt{Hz}$, and these of the gyroscope are set as $0.01°/h$ and $0.01°/\sqrt{h}$. The DVL scale factor error is set as 0.005. The update frequency of the IMU and DVL are set as $200Hz$ and $1Hz$, respectively.
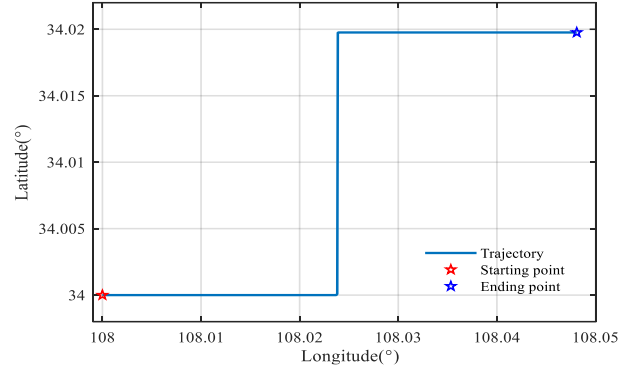
Fig. 4.   Vehicle trajectory.

In order to compare the performance of these methods, the noise with variance as (26) is generated:

$$R_k = \begin{cases} 0.3^2 & 700s \ \ to \ \ 800s \\ 2^2 & 1200s \ \ to \ \ 1300s \\ 0.1^2 & other \ \ time \end{cases} \qquad (26)$$

The initial model sets of IMMKF and IIMMKF are $\{0.1^2 \quad 0.3^2 \quad 0.5^2\}$. The velocity and position errors of the three algorithms are compared in Fig.5 and Fig.6. When the noise occurs, the conventional KF immediately diverges because of the incorrect measurement noise. Conversely, with more than one model, IMMKF and IIMMKF can avoid divergence to some extent. When the noise of 0.3 m/s occurs, the model sets of IMMKF and IIMMKF cover the true model appropriately and there is no filter divergence. However, when the noise of 3 m/s occurs, the model sets of IMMKF cannot cover the true model and filter diverges. With the adaptive model set and transition probability matrix, the model sets of IIMMKF can cover the true model in real time and the filter accuracy is improved.
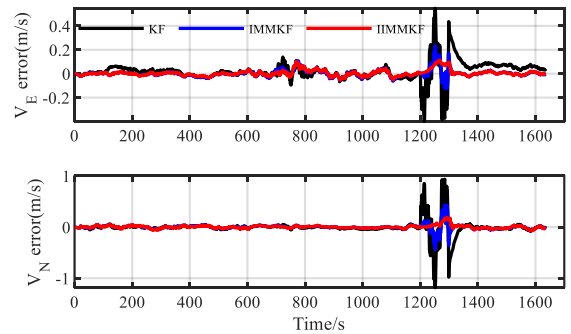
Fig. 5.   The velocity error of three methods.
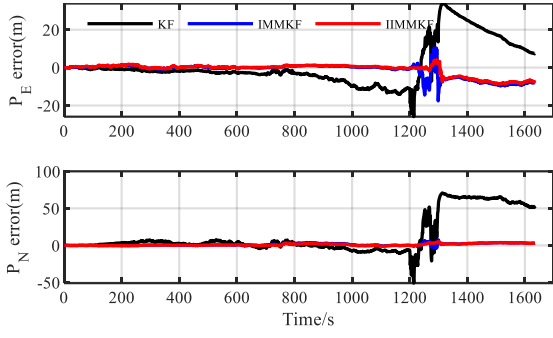
Fig. 6.   The position error of three methods.

Quantitative analysis can be done using the mean and root mean square (RMS) of the errors. The mean and RMS of the errors are shown in TABLE I and TABLE II. Comparing with the IMMKF method, the percentage improvement in the mean values of velocity and position errors are above 11%, 26%, 16%, and 14%. And the percentage improvement in the RMS values are above 20%, 46%, 2%, and 18%. It can be obviously seen that the proposed IIMMKF outperforms KF and IMMKF.

TABLE I.   THE MEAN OF THE VELOCITY AND POSITION ERRORS IN THREE METHODS

Parameters	Algorithm		
	KF	*IMMKF*	*IIMMKF*
V_E error (m/s)	0.046899115	0.018798682	0.016658681
V_N error (m/s)	0.053860774	0.027336851	0.020193287
P_E error (m)	7.713496469	2.215524458	1.850948948
P_N error (m)	16.85316961	1.353047544	1.151605375

TABLE II.   THE RMS OF THE VELOCITY AND POSITION ERRORS IN THREE METHODS

Parameters	Algorithm		
	KF	*IMMKF*	*IIMMKF*
V_E error (m/s)	0.079990697	0.031788761	0.025366122
V_N error (m/s)	0.149286228	0.056061583	0.030257747
P_E error (m)	0.055710953	0.014447561	0.014124397
P_N error (m)	11.24472785	3.89120056	3.173066311

B. *Vehicle Test*

The proposed method is testified in the land vehicle field test to predict the feasibility in the underwater environment. In the simulated INS/DVL navigation system of vehicle test, the inertial information is provided by the IMU and the velocity information of DVL is replaced by the PHINS developed by French firm IXBLU. PHINS provides the information of DVL by transforming its own velocity from navigation frame to body frame using the true attitude information. A computer is utilized to perform a series of navigation operations. The specifications of IMU and PHINS are listed in TABLE III.

TABLE III.   SPECIFICATIONS OF IMU AND PHINS

Sensor	Parameters	Accuracy	Rate
IMU	Gyroscope bias stability	$\leq 0.02°/h$	200Hz
	Gyroscope random walk	$\leq 0.005°/\sqrt{h}$	200Hz
	Accelerometer bias variation	$\pm 50ug$	200Hz
	Accelerometer output noise	$\leq 50ug/\sqrt{Hz}$	200Hz
PHINS	Attitude	$\leq 0.01°$	200Hz

Land trial is conducted near 31° 89'N, 118° 82'E, in the campus of Southeast University. A subset of about 1200s is used for evaluating the performance of the navigation system. The vehicle trajectory is shown in Fig. 7.
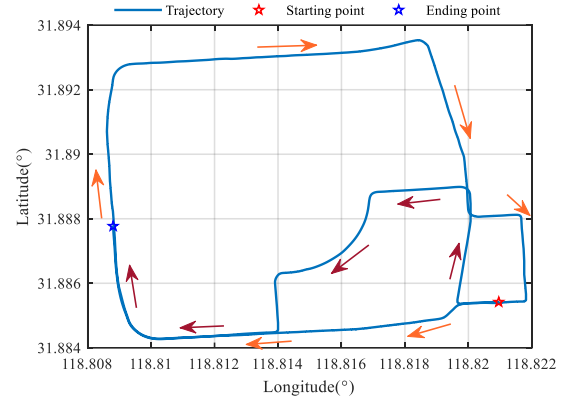
Fig. 7.   Vehicle trajectory.

In order to compare the performances of these algorithms, the noise with variance as (27) is generated to the INS/DVL navigation system.

$$R_k = \begin{cases} 0.3^2 & 500s \ to \ 600s \\ 2^2 & 1000s \ to \ 1100s \\ 0.1^2 & other \ time \end{cases} \quad (27)$$

The initial model sets of IMMKF and IIMMKF are $\left\{ 0.1^2 \quad 0.3^2 \quad 0.5^2 \right\}$ .The velocity and position errors are compared in Fig. 8 and Fig. 9. When the noise occurs, the conventional KF immediately diverges because of the incorrect measurement noise. Conversely, owing to more than one model, IMMKF and IIMMKF can avoid divergence to some extent. When the noise of 0.3 m/s occurs, the model sets of IMMKF and IIMMKF cover the true model appropriately and there is no filter divergence. However, when the noise of 3 m/s occurs, the model sets of IMMKF cannot cover the true model and filter diverges. With the adaptive model set and transition probability matrix, the model sets of IIMMKF can cover the true model in real time and the filter result is more accurate than others.

The mean and RMS of the errors are shown in TABLE IV and TABLE V. The contrast is especially sharp right at this moment. Comparing with the IMMKF method, the percentage improvement in the mean values of velocity and position errors are above 18%, 12%, 9%, and 1%. RMS values are improved up to 13%, 25%, 3%, and 16%.It can be obviously seen that the proposed IIMMKF outperforms KF and IMMKF.
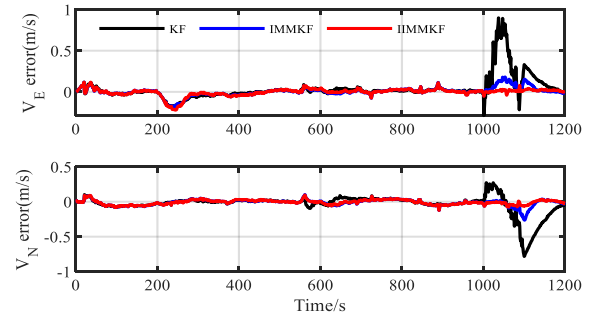
Fig. 8.   The velocity error of three methods.
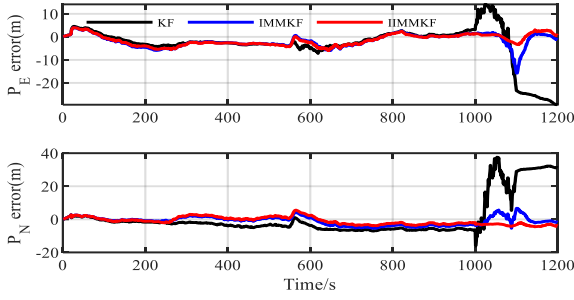
Fig. 9. The position error of three methods.

TABLE IV. THE MEAN OF THE VELOCITY AND POSITION ERRORS IN THREE METHODS

Parameters	Algorithm		
	KF	IMMKF	IIMMKF
V_E (m/s)	0.065513	0.03436	0.028091
V_N (m/s)	0.067981	0.031123	0.0272
P_E (m)	5.096753	2.573386	2.331776
P_N (m)	7.417376	2.186795	2.159385

TABLE V. THE RMS OF THE VELOCITY AND POSITION ERRORS IN THREE METHODS

Parameters	Algorithm		
	KF	IMMKF	IIMMKF
V_E (m/s)	0.144417779	0.052808418	0.045908055
V_N (m/s)	0.139146925	0.045094655	0.033797379
P_E (m)	0.065319514	0.032344143	0.031122649
P_N (m)	8.593295095	3.295615286	2.756178924

V. CONCLUSION

To realize the optimal estimation in the case of time-varying measurement noise and improve the accuracy of INS/DVL integrated navigation system, an IIMM algorithm is proposed in this paper. The model-probability-weighted-average-based adaptive model set is employed to adjust the model set to cover the true model in real time. Meanwhile, an adaptive transition probability matrix is introduced to accelerate the switch to the true model when the true model is not covered in the current model set. The IIMM algorithm enables the INS/DVL integrated navigation system to operate in uncertain, complex, and changing situations. Simulation and experiment show that the proposed algorithm can realize rapid model switch to cover the true model in real time and improve accuracy of thess INS/DVL integrated navigation system.

REFERENCES

[1] T. Zhang, J. Tang, S. Qin, and X. Wang, "Review of Navigation and Positioning of Deep-sea Manned Submersibles," Journal of Navigation, vol. 72, pp. 1021-1034, 2019.

[2] J. Tong, X. Xu, T. Zhang, Y. Li, Y. Yao, C. Weng, L. Hou, and L. Zhang, "A misalignment angle error calibration method of underwater acoustic array in strapdown inertial navigation system/ultrashort

[3] A. Tal, I. Klein and R. Katz, "Inertial Navigation System/Doppler Velocity Log (INS/DVL) Fusion with Partial DVL Measurements," Sensors, vol. 17, p. 415, 2017.

[4] Y. Yao, X. Xu and X. Xu, "An IMM-Aided ZUPT Methodology for an INS/DVL Integrated Navigation System," Sensors, vol. 17, p. 2030-, 2017.

[5] Y. Yang and T. Xu, "An Adaptive Kalman Filter Based on Sage Windowing Weights and Variance Components," Journal of Navigation, vol. 56, pp. 231-240, 2003.

[6] S. Akhlaghi, N. Zhou and Z. Huang, "Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation," in 2017 IEEE Power & Energy Society General Meeting, 2018.D. R. Salem, C. O'Driscoll, and G. Lachapelle, "Methodology for comparing two carrier phase tracking techniques," GPS Solutions, vol. 16, pp. 197–207, April 2012.

[7] Y. Liu, X. Fan, C. Lv, J. Wu, L. Li, and D. Ding, "An innovative information fusion method with adaptive Kalman filter for integrated INS/GPS navigation of autonomous vehicles," Mechanical Systems and Signal Processing, vol. 100, pp. 605-616, 2018.

[8] M. Narasimhappa, A. D. Mahindrakar, V. C. Guizilini, M. H. Terra, and S. L. Sabat, "An improved Sage Husa adaptive robust Kalman Filter for de-noising the MEMS IMU drift signal,", 2018, pp. 229-234.

[9] R. R. Pitre, V. P. Jilkov and X. R. Li, "A comparative study of multiple-model algorithms for maneuvering target tracking,". vol. 5809, 2005, pp. 549-560.

[10] L. Wang and H. T. Gao, "A variable structure multiple model filtering for SINS/DVL integrated solution," in Chinese Control Conference, T. Liu and Q. Zhao, Eds. NEW YORK: IEEE, 2017, pp. 6145-6150.

[11] Q. Xu, X. Li and C. Chan, "Enhancing Localization Accuracy of MEMS-INS/GPS/In-Vehicle Sensors Integration During GPS Outages," IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, vol. 67, pp. 1966-1978, 2018.

[12] Y. Li, X. Li, V. Havyarimana, D. Wang, and Z. Xiao, "Low-cost sensors aided vehicular position prediction with partial least squares regression during GPS outage," INTERNATIONAL JOURNAL OF EMBEDDED SYSTEMS, vol. 8, pp. 125-134, 2016.

[13] I. Hwang, C. E. Seah and S. Lee, "A Study on Stability of the Interacting Multiple Model Algorithm," IEEE TRANSACTIONS ON AUTOMATIC CONTROL, vol. 62, pp. 901-906, 2017.

[14] H. Song and Y. Choi, "Distributed multiple model extended information filter with unbiased mixing for satellite launch vehicle tracking," INTERNATIONAL JOURNAL OF DISTRIBUTED SENSOR NETWORKS, vol. 14, 2018.

[15] C. M. Huang and X. Wu, "Cooperative Vehicle Tracking Using Particle Filter Integrated with Interacting Multiple Models," in IEEE International Conference on Communications NEW YORK: IEEE, 2019.

[16] Y. Yao, X. Xu, Y. Li, and T. Zhang, "A Hybrid IMM Based INS/DVL Integration Solution for Underwater Vehicles," IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, vol. 68, pp. 5459-5470, 2019.

[17] Gao, W., Li, J., Zhou, G., and Q., "Adaptive Kalman Filtering with Recursive Noise Estimator for Integrated SINS/DVL Systems," Journal of Navigation, 2015.

[18] P. Liu, B. Wang, Z. Deng, and M. Fu, "A Correction Method for DVL Measurement Errors by Attitude Dynamics," IEEE Sensors Journal, vol. 17, pp. 4628-4638, 2017.

[19] Y. Liu, R. Song and R. Bucknall, "Intelligent Tracking of Moving Ships in Constrained Maritime Environments Using AIS," CYBERNETICS AND SYSTEMS, vol. 50, pp. 539-555, 2019.

[20] H. B., H. H., L. L., H. C., and Z. Z., "An Improved IMM Algorithm Based on STSRCKF for Maneuvering Target Tracking," IEEE Access, vol. 7, pp. 57795-57804, 2019-01-01 2019.

[21] M. D., D. B. and S. W., "The interacting multiple model algorithm based on adaptive Markov transition probability," in 2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), 2017, pp. 1-6.

[22] Z. Guo, C. Dong, Y. Cai, and Z. Yu, "Time-varing transition probablility base IMM-SRCKF algorithm for maneuvering target tracking," Systems Engineering and Electronics, vol. 37, pp. 24-30, 2015.

The Implementation and Comparison Between Kalman Filter-based and Vector Tracking Loops

Zhe Yan
Key Laboratory of Micro-Inertial Instrument and Advanced Navigation Technology, Ministry of Education School of Instrument Science and Engineering Southeast University
Nanjing, China
seuyanzhe@163.com

Xiyuan Chen*
Key Laboratory of Micro-Inertial Instrument and Advanced Navigation Technology, Ministry of Education School of Instrument Science and Engineering Southeast University
Nanjing, China
chxiyuan@seu.edu.cn

Xinhua Tang
Key Laboratory of Micro-Inertial Instrument and Advanced Navigation Technology, Ministry of Education School of Instrument Science and Engineering Southeast University
Nanjing, China
xinhuatanggnss@163.com

Xuefen Zhu
Key Laboratory of Micro-Inertial Instrument and Advanced Navigation Technology, Ministry of Education School of Instrument Science and Engineering Southeast University
Nanjing, China
zhuxuefen@seu.edu.cn

Abstract—A well-designed tracking loop is the core of a Global Navigation Satellite System (GNSS) receiver and plays a decisive role in estimating parameters and positioning. To solve the problem that classical scalar tracking may fail to track the received signal frequently in high-dynamic and weak signal environments, novel tracking loops based on optimal estimation are proposed by many researchers. Among them, Kalman-based and vector tracking are expected to have a better performance in improving the precision of measurements and positions. Considering their superiority and promising applications in deeply coupled GNSS/ inertial navigation system, the theoretical analysis and practical implementation are worthy of being further explored. In this paper, an EKF-based tracking loop which uses the outputs of coherent channels directly is proposed. And then, the implementation of a vector tracking frequency-locked and delay-locked loop is further proposed in detail. At last, the performance and superiority of the two methods are tested where a software-defined receiver is constructed and the aircraft data from a GNSS constellation simulator is used. Compared with scalar tracking, a significant improvement on positioning and measuring parameters is achieved by the two methods. And the vector tracking indicates a 42.9% and a 52.3% increase on the precision of positioning and pseudo-range measuring respectively.

Keywords—GNSS, tracking loops, Kalman filter, vector tracking

I. INTRODUCTION

The Global Navigation Satellite System (GNSS) has been widely used in various civil and military fields. However, the pursuit of better performance of the tracking loop in a GNSS receiver has never diminished among academic community.

Based on the classical control theory, a conventional scalar tracking loop is always a trade-off between accuracy and dynamic tolerance [1][2]. As a result, the performance of scalar loops may deteriorate dramatically when the receiver is in GNSS denied environments which most commonly include high dynamic and weak signal scenarios [3]. To maintain efficient tracking in these situations, Kalman-filter based

tracking loops have been proposed [4]. In a stand-along GNSS receiver, there are generally two kinds of architectures for signal tracking using Kalman filters according to different measurements. One of the structures uses the outputs of conventional scalar discriminators and is implemented by a linear Kalman filter. In the other one, the results of coherent accumulation are exploited and processed by a nonlinear Kalman filter[5]. The signal tracking loops based on Kalman filter have been studied and tested by many researchers with particular emphasis on PLL [5][6][7][8]. In these methods, the errors of scalar discriminators can be decreased and the limited linear range can be improved by nonlinear filters. However, the structures and implementations of the Kalman-based tracking loops are still diverse and difficult to be tuned. Recently, the equivalence between scalar tracking and Kalman-based tracking has been proved by many researchers [1][4][9], and the tuning criteria is available in reference [2] and [9] which provide a good opportunity for practical implementation.

Another way to implement the tracking loop using optimal estimation theory is vector tracking. Different from the scalar and Kalman-based tracking, vector tracking completes the signal tracking and the positioning simultaneously by a stand-along filter [10][11]. The feedback to construct the closed loop is computed using the spatial relationship between the receiver and visible satellites. Besides, the information from different channels are fused and fully utilized.

On the basis of the tracking loop realized by optimal estimations, inertial navigation system (INS) can be added to constitute a GNSS/INS deeply coupled system which performs a much more robust tracking loop. In these INS-assisted approaches, Kalman-based tracking is the basis of federated schemes of the deeply coupled system [3] while vector tracking the centralized [12]. For further researches and studies on robust tracking loops, we need to implement and analyze the Kalman-based and vector tracking in detail. So the detailed implementation of Kalman-based and vector tracking loops are proposed herein, and their performance is tested and compared using high-dynamic aircraft data.

---

*Xiyuan Chen is the corresponding author. (e-mail: chxiyuan@seu.edu.cn).

II. SCALAR TRACKING LOOPS

A. The Basic Principle of GNSS Tracking Loops

To carry out the calculation of the position and velocity of a GNSS receiver at a certain time (Position, Velocity, and Time, PVT), a few procedures of signal processing are needed, including the down-conversion of radio frequency, acquisition, tracking, bit synchronization, and frame synchronization. Among them, signal tracking is the core part in a GNSS receiver, which exert a great influence on the performance of the receiver. The main job of the tracking loops is to estimate the Doppler frequency f_d, initial carrier-phase φ and code-delay τ of the received signal accurately and dynamically, maintaining the local replica of satellite signals the same with the received ones.

Suppose k and T represent the sequence number and the interval of sampling respectively, the received satellite signal $r(k)$ with random noise n and intermediate frequency f_{IF} can be discretized as

$$r(k) = A \cdot C(kT - \tau) \cdot \cos\left(2\pi kT\left(f_{IF} + f_d\right) + \varphi\right) + n(kT), \quad (1)$$

where A and C denote the amplitude of $r(k)$ and the sequence of pseudo-code respectively.

B. Scalar Tracking Loops with Conventional Discriminators

Scalar Tracking loops usually refer to the frequency-locked loops (FLL), phase-locked loops (PLL), delay-locked loops (DLL), and their combinations which estimate Doppler frequency f_d, initial carrier-phases φ and code-delays τ respectively. A typical architecture of the scalar tracking can be seen in Fig.1.
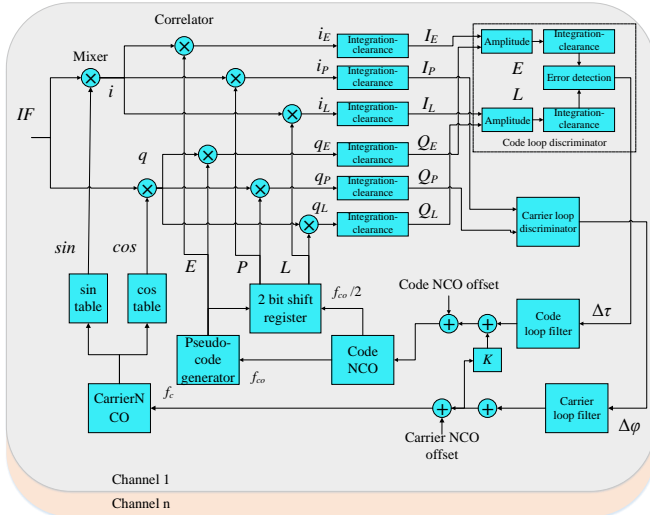
Fig. 1.   The typical scalar tracking loop with discriminators.

As illustrated in Fig.1, scalar tracking loops include the independently performed carrier loop and code loop which consist of correlation channels, discriminators, loop filters and numerically controlled oscillators (NCO). Discriminators are used to compute the differences between the parameters of replica signal and the received signal, then the outputs will be filtered by loop filters to calculate the feedback of the closed loops. Generally, the error of f_d obtained from a FLL can be expressed by

$$\Delta f = \frac{cross \times sign(dot)}{T_{coh}}, \quad (2)$$

where T_{coh} is the interval of coherent accumulation, and

$$cross = I_{P1} \times Q_{P1} - I_{P2} \times Q_{P2}; \quad (3)$$

$$dot = I_{P1} \times I_{P2} + Q_{P1} \times Q_{P2}. \quad (4)$$

Suppose I_P/Q_P, I_E/Q_E, and I_L/Q_L are the coherent accumulations of prompt, early and late branches respectively, and their code phase interval with each other is half a code chip. The I and Q denote in-phase and quadrature-phase. The error of φ and τ estimated by the discriminators can be given by

$$\Delta \varphi = \arctan \frac{Q_P}{I_P}; \quad (5)$$

$$\Delta \tau = \frac{\sqrt{I_E^2 + Q_E^2} - \sqrt{I_L^2 + Q_L^2}}{\sqrt{I_E^2 + Q_E^2} + \sqrt{I_L^2 + Q_L^2}}. \quad (6)$$

III. EKF-BASED TRACKING LOOPS

In conventional scalar tracking loops, the pull-in ranges and nonlinear characteristics of the discriminators limit their scope of application. As a result, carrier-smoothing or a Kalman filter are usually used in the calculation of PVT. Another way to improve the performance of discriminators is introducing the theory of optimal estimation into the tracking loops. For example, the outputs of the discriminators can be used as the observations of a Kalman filter so that the errors of these outputs can be narrowed down. The equivalence of the conventional and the Kalman based tracking loops can be seen in our previous work [9].

In this section, a tracking loop based on an extended Kalman filter (EKF) is proposed in which conventional discriminators are replaced with an EKF. The scheme diagram of the EKF-based tracking loop proposed can be seen in Fig.2.
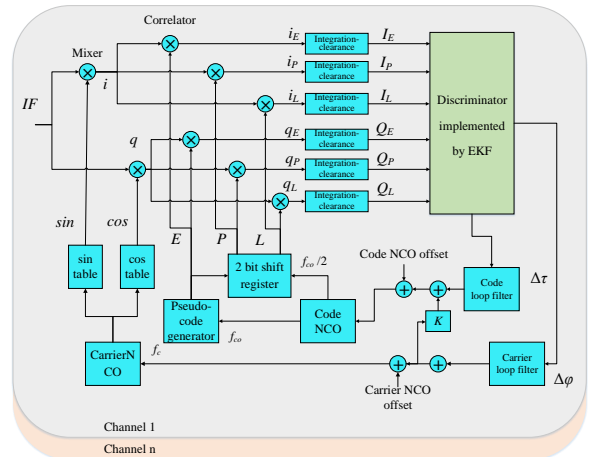
Fig. 2.   EKF-based tracking loop without conventional discriminators.

The sate parameters of the filter are given by

$$\mathbf{X} = \begin{bmatrix} A & \Delta\tau & \Delta\varphi & \Delta f_d & \Delta f_a \end{bmatrix}^T, \quad (7)$$

where f_a is the changing rate of the Doppler frequency f_d. The state equation can be described as

$$\mathbf{X}_{k+1} = f\left(\mathbf{X}_k, k\right) + \mathbf{W}_k . \tag{8}$$

The transmission function of the states and the noise of the system are given by

$$f\left(\mathbf{X}_k, k\right) = \begin{bmatrix} A_k \\ \dfrac{\beta T_{coh}}{2\pi} \Delta f_k \\ \Delta \varphi_k + T_{coh}\Delta f_k + \dfrac{1}{2}T_{coh}^2 \Delta f_{a,k} \\ \Delta f_k + T_{coh}\Delta f_{a,k} \\ \Delta f_{a,k} \end{bmatrix} ; \tag{9}$$

$$\mathbf{W}_k = \begin{bmatrix} \mathbf{W}_A & \mathbf{W}_\tau & \mathbf{W}_\varphi & \mathbf{W}_f & \mathbf{W}_a \end{bmatrix}, \tag{10}$$

where β is the ratio between the frequencies of the carrier and the pseudo-code. This coefficient can be given by $\beta = 1575.42MHz/1.023MHz$ for GPS L1. The system noise $\mathbf{W}_k$ includes the process noise of ΔA, $\Delta \tau$, $\Delta \varphi$, Δf, and Δf_a.

The observation equation can be described as

$$\mathbf{Z}_k = h\left(\mathbf{X}_k, k\right) + \mathbf{V}_k . \tag{11}$$

The observations, observation matrix and measurement noise are given by following equations respectively.

$$\mathbf{Z}_k = \begin{bmatrix} I_E & Q_E & I_P & Q_P & I_L & Q_L \end{bmatrix}_k^T ; \tag{12}$$

$$h\left(\mathbf{X}_k, k\right) =$$
$$\begin{bmatrix} \dfrac{NA_k}{2}R\left(\Delta\tau_k + 0.5\right)\operatorname{sinc}\left(\alpha\right)\sin\left(\chi\right) \\ \dfrac{NA_k}{2}R\left(\Delta\tau_k + 0.5\right)\operatorname{sinc}\left(\alpha\right)\cos\left(\chi\right) \\ \dfrac{NA_k}{2}R\left(\Delta\tau_k\right)\operatorname{sinc}\left(\alpha\right)\sin\left(\chi\right) \\ \dfrac{NA_k}{2}R\left(\Delta\tau_k\right)\operatorname{sinc}\left(\alpha\right)\cos\left(\chi\right) \\ \dfrac{NA_k}{2}R\left(\Delta\tau_k - 0.5\right)\operatorname{sinc}\left(\alpha\right)\sin\left(\chi\right) \\ \dfrac{NA_k}{2}R\left(\Delta\tau_k - 0.5\right)\operatorname{sinc}\left(\alpha\right)\cos\left(\chi\alpha\right) \end{bmatrix} ; \tag{13}$$

$$\mathbf{V}_k = \begin{bmatrix} \eta_{IE} & \eta_{QE} & \eta_{IP} & \eta_{QP} & \eta_{IL} & \eta_{QL} \end{bmatrix}_k^T , \tag{14}$$

where R and N are the self-correlation function and the number of the samples in an accumulation interval respectively. η denotes zero mean white noise. α and χ are described by

$$\alpha = \Delta\varphi_k + \dfrac{T_{coh}}{2}\Delta f_k + \dfrac{T_{coh}^2}{6}\Delta f_{a,k} ; \tag{15}$$

$$\chi = \dfrac{\Delta f_k}{2}T_{coh} . \tag{16}$$

Because the tracking loop based on EKF is a closed loop with the feedback to NCOs, the one-step prediction $\hat{\mathbf{X}}_{k/k-1}$ is assumed to be zero after the NCOs' being adjusted by

feedback. So the one-step prediction and the state estimation need to be modified by

$$\hat{\mathbf{X}}_{k/k-1} = 0 ; \tag{17}$$

$$\hat{\mathbf{X}}_k = \mathbf{K}_k \mathbf{Z}_k , \tag{18}$$

where $\hat{\mathbf{X}}_k$ and $\mathbf{K}_k$ represent the estimation of the state and the gain of the filter respectively.

IV. VECTOR TRACKING LOOPS

A. Introduction of Vector Tracking

Vector tracking is drawing a growing interest among academic community and is proposed to solve the problem that the conventional scalar-tracking loop is so sensitive to noise and dynamics that it may fail to lock the signal in multipath, weak signal, strong jamming, and high dynamic environments.

By optimal estimation theory, the information from different channels and different satellites is fused together and aided by each other. The other key characteristic of vector tracking is that the PVT is obtained from the Kalman filter directly, and then the feedback of the loop is constructed using the line-of-sight (LOS) vectors between the receiver and the satellites tracked. This method can make full use of the spatial relationship between different signals and is expected to provide a better performance in challenged and GNSS denied scenarios. As a result, vector tracking is now considered as an important foundation of deeply coupled system using GNSS and inertial navigation system [12]. But it is important to note that though the tracking channels are capable of assisting with each other, they may be polluted by the deteriorative channels. A VDFLL composed of the vector tracking FLL (VDLL) and vector tracking FLL (VFLL) can be seen in Fig.3.
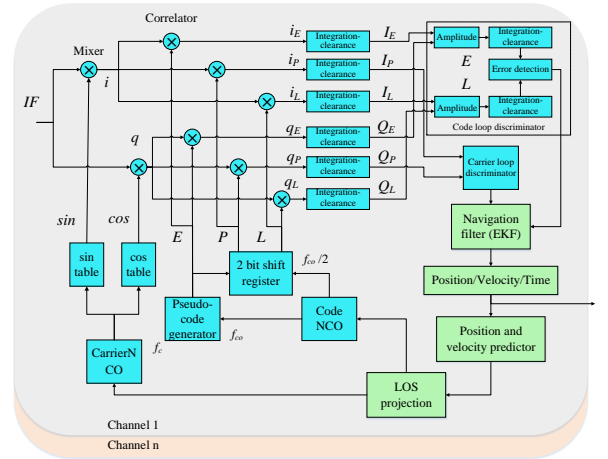
Fig. 3.   The vector tracking loop with VDFLL.

B. Design of the Navigation Filter

In our research, the navigation filter is performed by an EKF whose state equation is given by

$$\begin{bmatrix} \Delta\mathbf{P}_{k+1} \\ \Delta\mathbf{V}_{k+1} \\ t_{b,k+1} \\ t_{d,k+1} \end{bmatrix} = \mathbf{\Phi}_{k/k+1} \begin{bmatrix} \Delta\mathbf{P}_k \\ \Delta\mathbf{V}_k \\ t_{b,k} \\ t_{d,k} \end{bmatrix} + \mathbf{W}_k ; \tag{19}$$

$$\Delta\mathbf{P}_k = \begin{bmatrix} P_{x,k} - \hat{P}_{x,k} & P_{y,k} - \hat{P}_{y,k} & P_{z,k} - \hat{P}_{z,k} \end{bmatrix}^T ; \quad (20)$$

$$\Delta\mathbf{V}_k = \begin{bmatrix} V_{x,k} - \hat{V}_{x,k} & V_{y,k} - \hat{V}_{y,k} & V_{z,k} - \hat{V}_{z,k} \end{bmatrix}^T ; \quad (21)$$

$$\mathbf{\Phi}_{k/k+1} = \begin{bmatrix} 1 & 0 & 0 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & T & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & T & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (22)$$

where $\Delta\mathbf{P}$ and $\Delta\mathbf{V}$ are the errors of position and velocity in three directions. t_b and t_d denote the bias and drift of the clock respectively.

The observations are defined by the outputs of discriminators as follows

$$\mathbf{Z}_k = \begin{bmatrix} Z^1_{code,k} & \text{L} & Z^j_{code,k} & Z^1_{carrier,k} & \text{L} & Z^j_{carrier,k} \end{bmatrix}^T \quad (23)$$

where $Z^j_{code,k}$ is the code-delay $\Delta\tau$ converted from chips to meters, and $Z^j_{carrier,k}$ is the Doppler frequency Δf converted from Hertz to meter per second. j means the jth channel. Then, the observation equation can be described as

$$\mathbf{Z}_k = \mathbf{H}_k \Delta\mathbf{X}_k + \mathbf{V}_k ; \quad (24)$$

$$\mathbf{H}_k = \begin{bmatrix} a^1_x & a^1_y & a^1_z & 0 & 0 & 0 & -1 & 0 \\ \text{M} & \text{M} & \text{M} & \text{M} & \text{M} & \text{M} & \text{M} & \text{M} \\ a^j_x & a^j_y & a^j_z & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & a^1_x & a^1_y & a^1_z & 0 & -1 \\ \text{M} & \text{M} & \text{M} & \text{M} & \text{M} & \text{M} & \text{M} & \text{M} \\ 0 & 0 & 0 & a^j_x & a^j_y & a^j_z & 0 & -1 \end{bmatrix}, \quad (25)$$

where $\begin{bmatrix} a^j_x & a^j_y & a^j_z \end{bmatrix}^T$ is the LOS vector from the receiver to the satellite of the jth channel.

V. EXPERIMENT AND DISCUSSION

In order to fully verify the excellent performance of the EKF based tracking loop and VDFLL, experiments using a constellation simulator and a soft-ware defined receiver were designed.
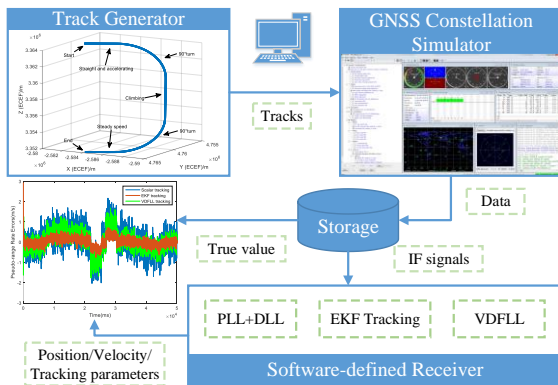
Fig. 4. Diagram of the experiment methods and procedures.

Experiments and comparisons were carried out using aircraft data that was sampled from a GNSS constellation simulator as described in Fig. 4. The maneuvers of the aircraft are described in Table I.

TABLE I. MANEUVERS OF THE SIMULATED AIRCRAFT

Time/s	Motions of the Aircraft
0-5	Accelerating from 200 to 600m/s
5-23	A 90°change in heading with steady speed, and the lateral acceleration is 6g.
23-28	Steady climbing in height with a change of 1000m, and lateral acceleration is 6g.
28-46	A 90°change in heading with steady speed, and the lateral acceleration is 6g.
46-50	Uniform motion with 600m/s.

The initial position is [32.03°N, 118.47°E, 63m], and the initial speed is 200m/s. The trajectory of the aircraft is described in Fig. 5 with earth-centered and earth-fixed (ECEF) coordinate frame. For GNSS signal, GPS L1 was chosen to simulate 9.548MHz intermediate-frequency (IF) signals with a sampling frequency of 38.192MHz. The simulation began from 2020-5-15 12:00:00. In this experiment, a 2rd PLL and 2rd DLL were used, where the bandwidths of DLL and PLL are 2 Hz and 25 Hz respectively.
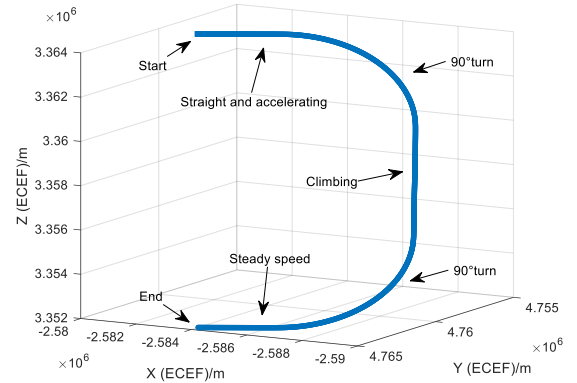
Fig. 5. The trajectory of the simulated aircraft.

The positioning errors of scalar, EKF and VDFLL tracking in three directions are described by Fig. 6 which indicates the best positioning accuracy of VDFLL and the worst of conventional scalar tracking. As shown in Fig.6, the curve of VDFLL is closer to zero and fluctuates in a smaller range. And the accuracy of the EKF falls somewhere in between.

(a)
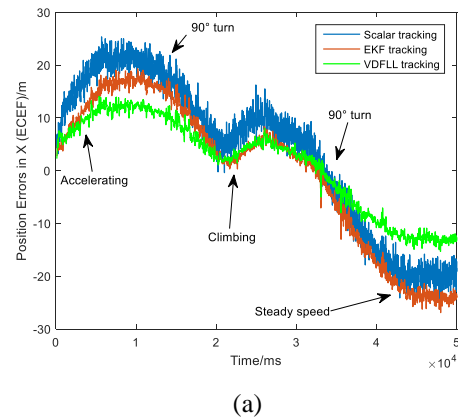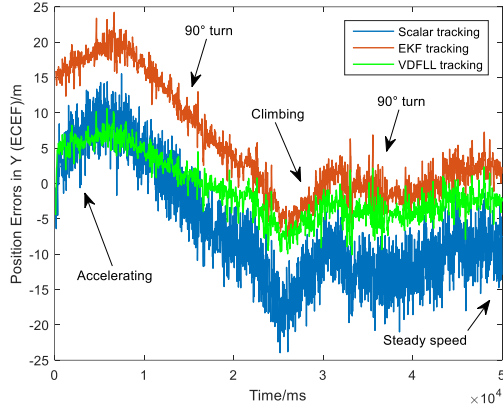
(b)
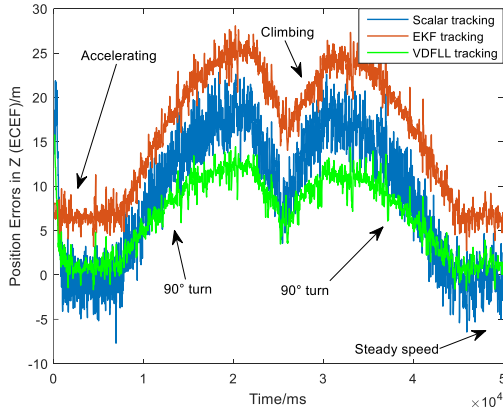
(c)

Fig. 6. Positioning errors in three directions: (a) X; (b) Y; (c) Z axis of ECEF.

The standard deviations of the positioning errors are summarized in Table II. As we can see from the table, an obvious improvement on positioning precision was achieved by EKF tracking loop and vector tracking loop. Compared with scalar tracking, the decrease on standard deviations are 3.7% and 42.9% respectively in Z direction. For EKF tracking, the errors of discriminators are narrowed down using EKF so that the measurements for PVT calculation, namely pseudo-range and pseudo-range rate can be improved.

TABLE II.   STANDARD DEVIATIONS OF THE POSITIONING ERRORS

Methods	Standard deviations (m)		
	PLL+DLL	EKF tracking	VDFLL
X	14.38	13.96	8.73
Y	8.26	7.96	4.37
Z	7.35	7.08	4.20

For vector tracking, the improvement on positioning accuracy mainly benefits from the PVT calculation using a Kalman filter. The superiority of Kalman based PVT calculation (KF-PVT) compared with the least-squares PVT (LS-PVT) has been proved [2]. Another, the measurements of PVT are also improved because the feedback constructed by the positions obtained from KF-PVT seems to be more accurate than that by original outputs of discriminators, but

this also depends on the configuration and the performance of the loop filter.

The improvements on one of the PVT measurements, the pseudo-ranges, are shown in Fig.7.
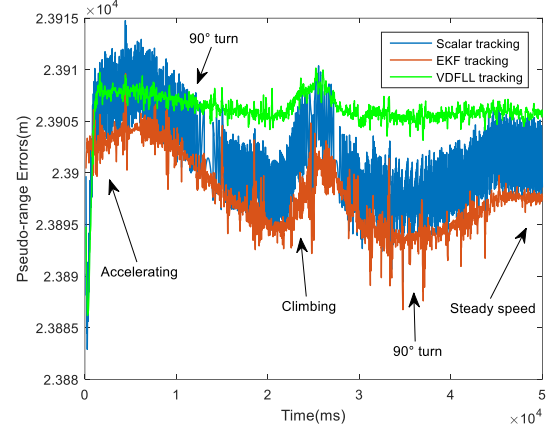
Fig. 7. Pseudo-range errors of the three methods (PRN 2).

Fig. 7 illustrate that both EKF tracking and VDFLL are more stable in the precision of pseudo-range, because the amplitude of oscillation is much smaller. This phenomenon is also indicated in Fig.8 where the errors of pseudo-ranges corrected by clock bias are described. The standard deviations of pseudo-range errors are summarized in Table III. For the precision of not corrected pseudo-range, the improvements are 17.3% and 52.3% for EKF and VDFLL tracking respectively.
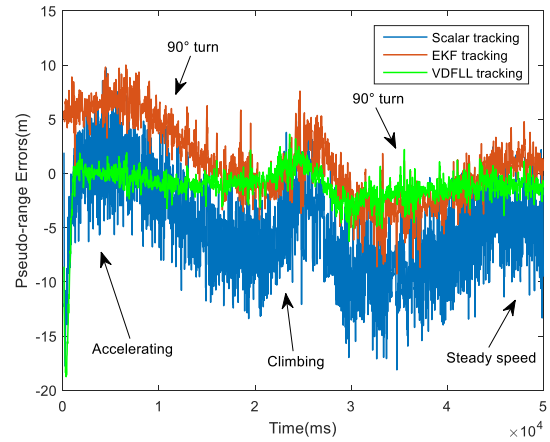
Fig. 8. Corrected pseudo-range errors of the three methods (PRN 2).

TABLE III.   STANDARD DEVIATIONS OF THE PSEUDO-RANGE ERRORS

Methods	Standard deviations (m)		
	PLL+DLL	EKF tracking	VDFLL
Not corrected	4.40	3.64	2.10
Corrected	4.42	3.54	1.95

It can be concluded that the positioning performance of all the three methods are sensitive to the maneuvers of the aircraft, while the pseudo-range measurements of VDFLL are not as sensitive as others. The improvements on pseudo-range rate are shown in Fig. 9. And the standard deviations of the errors of pseudo-range rate are summarized by Table IV.

49
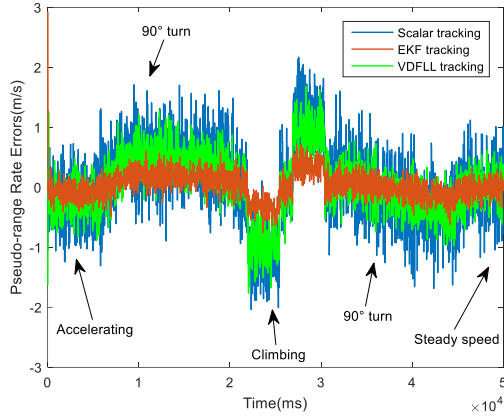
Fig. 9. Pseudo-range rate errors of the three methods (PRN 2).

TABLE IV.   STANDARD DEVIATIONS OF THE PSEUDO-RANGE RATE
ERRORS

Methods	Standard deviations (m/s)		
	PLL+DLL	**EKF tracking**	**VDFLL**
Pseudo-range rate	0.69	0.22	0.51

As we can see from Fig. 9 and Table IV, EKF has the most significant improvement on the measurement of pseudo-range rate (68.1% decline on standard deviation). Because the precision of FLL is generally much worse than PLL, the performance of VFLL is not as good as EKF based PLL. But VFLL still achieved a better performance than conventional scalar PLL (26.1% decline on standard deviation). The pseudo-range rates corrected by clock drift cannot be offered for the reason that the clock drift is not estimated by LS-PVT which is used by scalar tracking and EKF tracking loops designed in this paper.

VI. CONCLUSION

Based on the theory of optimal estimation, Kalman-based and vector tacking loops in a GNSS receiver are introduced in this paper. Then, an EKF-based PLL/DLL tracking loop and an EKF-based VDFLL are implemented on the platform of a soft-ware defined GPS L1 receiver. By the experiment and testing using the aircraft data simulated by the GNSS constellation simulator, the results indicate that VDFLL has the best performance on positioning and the improvement on pseudo-range measurements, while the EKF tracking falls between VDFLL and PLL/DLL. But the EKF tracking has the most significant improvement on pseudo-range rate measurements. Besides, the pseudo-range measurements of VDFLL are less sensitive to maneuvers than those of another two methods.

REFERENCES

[1] X. Niu, B. Li, N.I. Ziedan, et al, "Analytical and simulation-based comparison between traditional and Kalman filter-based phase-locked loops," *GPS Solutions*, vol. 21, pp. 123–135, January 2017.

[2] X. Tang, X. Chen, Z. Pei and P. Wang, "The Explicit Tuning Investigation and Validation of a Full Kalman Filter-Based Tracking Loop in GNSS Receivers," *IEEE Access*, vol. 7, pp. 111487-111498, July 2019.

[3] E. J. Ohlmeyer, "Analysis of an ultra-tightly coupled GPS/INS system in jamming,". Proceedings of the 2006 IEEE/ION Position, Location, & Navigation Symposium, San Diego, CA, USA, April 2006, pp. 44–53.

[4] R. Jiang, K. Wang and S. Liu, et al, " Performance analysis of a Kalman filter carrier phase tracking loop," *GPS Solutions*, vol. 21, pp. 551-559, April 2017.

[5] X. Tang, G. Falco, E. Falletti and L. L. Presti, "Practical implementation and performance assessment of an Extended Kalman Filter-based signal tracking loop," 2013 International Conference on Localization and GNSS (ICL-GNSS), Turin, Italy, June 2013, pp. 1-6.

[6] M. Petovello and G.Lachapelle, "Comparison of vector-based software receiver implementations with application to ultra-tight GPS/INS integration," in ION GNSS 2006, Fort Worth TX, 26-29, September 2006.

[7] D. R. Salem, C. O'Driscoll, and G. Lachapelle, "Methodology for comparing two carrier phase tracking techniques," *GPS Solutions*, vol. 16, pp. 197–207, April 2012.

[8] F. Macchi, "Development and testing of an L1 combined GPS-Galileo software receiver," Ph.D. dissertation, University of Calgary, January 2010.

[9] X. Tang, G. Falco, E. Falletti and L. L. Presti, "Theoretical analysis and tuning criteria of the Kalman filter-based tracking loop," *GPS Solutions*, vol. 19, no. 3, pp. 489-503, July 2015.

[10] M. Lashley, D. M. Bevly and J. Y. Hung, "Performance analysis of vector tracking algorithms for weak GPS signals in high dynamics," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 4, pp. 661-673, August 2009.

[11] Y. Ng and G. X. Gao, "GNSS multireceiver vector tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 5, pp. 2583-2593, October 2017.

[12] Z. Yan, X. Chen, X. Tang, "A novel linear model based on code approximation for GNSS/INS ultra-tight integration system," *Sensors* vol. 20, no. 11, pp.1-20, June 2020.

Graph Theory Based Localization of Wireless Sensor Networks for Radio Irregularity Cases

Xiaofeng Ma
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
buaamxf@buaa.edu.cn

Ning Yu*
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
nyu@buaa.edu.cn

Tianle Zhou
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
zpulse@buaa.edu.cn

Renjian Feng
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
rjfeng@buaa.edu.cn

Yinfeng Wu
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
yfwu@buaa.edu.cn

Abstract—**In the aerospace applications, the cases of radio irregularity, obstacle, asymmetrical distribution are disadvantageous for node localization of wireless sensor networks. In this paper, the related principles and methods of Graph Theory is introduced to minimize the localization error caused by radio irregularity. The model of asymmetry communication is established based on Graph Theory. Through the analysis to communication path between nodes in directed network, solution schemes are put forward for locating nodes in case of asymmetric communication. The localization performance is compared in different scenes and conditions. For node localization with radio irregularity, the accuracy can be improved 10% by our solutions when the network topology is regular. While the network topology is irregular, the localization accuracy can also be improved when strongly connected graph comes into being.**

Keywords—wireless sensor networks, localization, graph theory, radio irregularity, asymmetry communication

I. Introduction

Wireless sensor networks (WSN) are placed in the demand area and used to perceive the status of each position and communicate. They are self-powered by the on-board batteries. The application prospect of this technology in aerospace field is very broad. It can be used to monitor the state of spacecraft and other equipment[1], environment monitoring of the moon[2], Mars and other exploration planets[3, 4], physical condition monitoring of astronauts[5] and localization of unmanned exploration equipment[6, 7], etc.

Obviously, the exact location of the node needs to be known, otherwise the detected information will be meaningless, and there is a higher requirement for the location accuracy of the nodes in WSN in the aerospace field. However, due to the huge number of WSN nodes, it is unrealistic to accurately measure the location of nodes manually, and the GPS cannot be used in the space environment. At present, there are many algorithms to calculate the unknown node location by a small number of anchor nodes, they are fundamentally important for the applications and the location-based network protocol of WSN as the supporting technology of WSN.

As shown in Fig.1, the radio range of the sensor node is always irregular since the influence of anisotropic path loss, inconsistent transmit power and obstacles caused by factors such as rugged terrain in the explored region of the planets. It

is disadvantage to node localization, the MAC layer functions, the topology control, and network protocol etc. This is unfavorable to the environment monitoring, target tracking and other functions of sensor networks.

To solve the problem of asymmetry communication brought by radio irregularity, many scholars have studied it from different aspects. [8] studied the radio irregularity from two aspects: space and time. For the irregularity in space, they select the neighbors on the rim of the irregular radio coverage region as the forwarders. For the dynamic of the radio coverage in time, an on-demand, stateless strategy is presented. [9] integrated Fuzzy Logic and Extreme Learning Machines techniques to compensate for the effects of irregular topology.
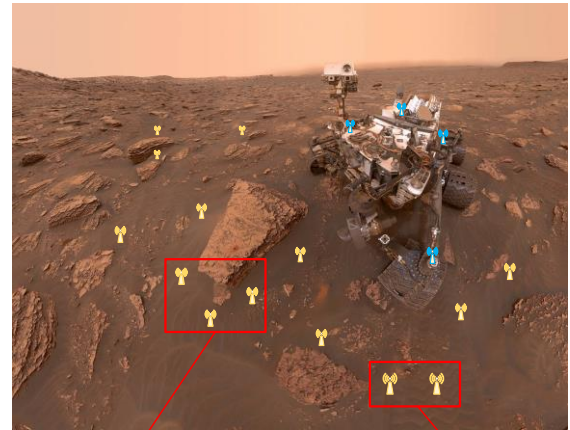
Image Credit: NASA/JPL-Caltech
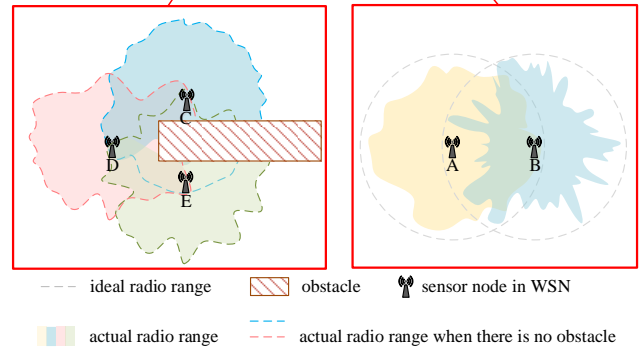
Fig. 1. Information propagation denoted by directed graph.

---

[10] considered the three dimensional aspects and non-isotropic radio model along with radio irregularity, then proposed a three dimensional radio irregularity model. [11] proposed a mobile beacon based range free localization method for WSN, which is based on analytical geometry of an arc. In this scheme, Cramers rule is used, where the intersection point of two perpendicular bisectors of the chords is taken as the estimated location of the sensor node. [12] proposed a Compressive Family-based Efficient Trust Routing Protocol by dividing the network into various clusters and then split up clusters into sub-clusters. Further, each sub-cluster is separated into various families and each family is allocated with a family head. [13] segmented the irregular coverage region into several regular bows and triangles by geometry decomposition approach, so that areas can be calculated conveniently.

The relevant principles and methods of Graph Theory are introduced to solve the problems in more and more research field of the WSN. [14] introduced a graph based coverage and connectivity technique that is aimed to attain better optimization of coverage and connectivity issues in large scale WSN considering the issues of energy depletion factor. [15] proposed a graph-theoretic distributed protocol to detect simultaneously the faults and cuts in the WSN. [16] proposed a new Graph Theory based methodology by which it identifies the optimum routing path from a source to destination so that there will be a minimum number of packet drop while it gains maximum throughput. [17] proposed a novel and efficient clustering called Clustering using Eigen Values with the increased lifetime of the sensor nodes using the spectral Graph Theory.

In this paper, applying the core idea of Graph Theory, we present a solution to reduce localization error caused by radio irregularity. In Section II, the relevant conceptions and methods of the Graph Theory are introduced as the theoretical basis of the localization in WSN under radio irregularity. The model of radio irregularity and directed communication is established. The solution of localization with radio irregularity is presented. In Section III, the localization performance is compared in different scenes and conditions. Finally, the conclusions are discussed in section IV.

II. LOCALIZATION WITH THE RADIO IRREGULARITY

A. The Radio Irregularity Model

In most existing research of localization in WSN, the radio range of the sensor node is assumed as a perfect round, but it is not actually due to the factors such as influence of anisotropic path loss, propagation energy and rock block on the surface of the planet. We use radio irregularity model to make sure that the research accords with the actual situation.
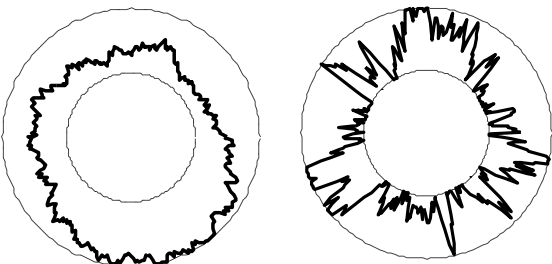
As shown in Fig. 2, it is assumed that the radio range of sensor node has an upper bound and a lower bound, i.e. the communication amplitude is different in different directions. If the distance between nodes is beyond the upper bound, the nodes cannot communicate with each other. On the contrary, the nodes can communicate with the neighbor nodes within the lower bound of the radio range. If the distance between two nodes is between the lower bound and the upper bound, there are three situations: *1) symmetric communication, 2) asymmetric communication,* and *3) no communication.*

The radio irregularity can be described by the degree of irregularity (DOI) which is defined as the maximum radio range variation per unit degree change in the direction of radio propagation [18]. If the DOI is zero, i.e. the upper and lower bounds coincide, the radio model is a perfect round and the intensity of the communication amplitude is the same in all directions. Fig. 2 shows the radio range in simulation with DOI is 0.05 and 0.2 respectively.

B. Sensor Network Model Based on Digraph

Radio irregularity causes the asymmetric communication between the neighbor nodes. On some direction, the node can send the information to its neighbor node but cannot receive the massage from the neighbor as the transmit power of the neighbor is weaker. This is the unidirectional (asymmetric) communication between two nodes that can be denoted by the Digraph in Graph Theory.

Definition 1: *Digraph.* The digraph G consists of two sets V and E, denoted as $G=(V, E)$. Where V is the finite nonempty set of the vertexes and E is the finite set of the edges between two vertexes in V. Each edge in G is directional.

Definition 2: *Degree of the vertex.* The degree of the vertex v, write as $deg(v)$, is the number of the vertexes with that v can connect.

As shown in Fig. 3, the sensor nodes and the communication between nodes in WSN can be expressed by the vertex and directional edge between vertexes in digraph, respectively. The arrow in Fig. 3 is the direction in which the information is transferred. The node at the end-point of the arrow can receive the information sent by the node at the start-point of the arrow.

In Graph Theory, the number of edges whose end-point is v is the indegree of v, denoted as $ID(v)$, and the number of the edges whose start-point is v is the outdegree of v, denoted as $OD(v)$. For example, the $ID(v)$ and the $OD(v)$ of node v in Fig. 3 are 1 and 2, respectively. The connectivity of v is defined as

Fig. 2. Radio irregularity model.
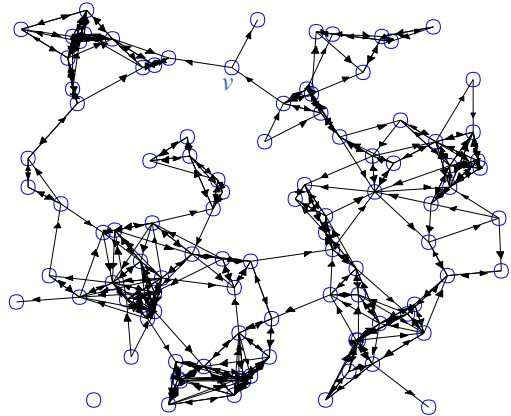
Fig. 3. Information propagation denoted by directed graph.

the summation of $ID(v)$ and $OD(v)$, write as $D(v)=ID(v)+OD(v)$. Consequently, the $D(v)$ of v in Fig. 3 is 3. The (1) shows the relationship of the node number n, the edge number e, and the connectivity $D(v)$.

$$e = \frac{1}{2}\sum_{i=1}^{n} D(v_i) \qquad (1)$$

Therefore, the average connectivity of the networks is $2e/n$.

The difference between $OD(v)$ and $ID(v)$ can be used as the signal strength intensity (SSI) of node v compared with the neighbor nodes, namely $SSI(v)= OD(v)–ID(v)$. The larger the $SSI(v)$, the stronger the signal of node v compared with neighbor nodes is. The value of SSI also shows the asymmetric information propagation.

C. Analysis of Communication Path between Nodes

The asymmetric communication in WSN causes the increase of the localization error. We devoted to that how to perform the node localization better under this situation.

The localization error of DV-distance algorithm [19] with asymmetric communication is analyzed through the simulation. In our simulation, two hundred senor nodes are randomly deployed in a 300×300 region. The upper bound of the radio irregularity is 53 and the lower bound is 27, the DOI=0.2. The result is shown in Fig. 6. The localization error changes between 27% and 35% with the increase of the anchor proportion. The factors that induce the localization error contain the asymmetric communication between nodes except the deficiency of the localization algorithm. The solutions are discussed for the asymmetric communication bellow.

Definition 3: *Path*. If there is a sequence of vertexes v_p, v_{i1}, $v_{i2}, \ldots , v_{im}$, v_q and $<v_p, v_{i1}>$, $<v_{i1}, v_{i2}>$, $\ldots$, $<v_{im}, v_q>$ are all belong to $E(G)$. There is a path from v_p to v_q.

In the undirected (symmetric) graph, if there is a path from v_p to v_q, there must be a path on opposite direction. In Digraph, there are three situations between v_p and v_q: 1) there are two opposite paths through the same nodes, 2) there are two opposite paths through different nodes, 3) there is only one or no path. The Adjacency Matrix and the shortest path in the Graph Theory are introduced to distinguish the paths.

Definition 4: *Adjacency Matrix of the Graph*. Set $G=(V, E)$, G is the graph with n vertexes. The Adjacency Matrix A of G is n order square matrix:

$$A[i,j] = \begin{cases} w_{ij} & if < v_i, v_j > \in E(G) \\ \infty & if < v_i, v_j > \notin E(G) \end{cases} \qquad (2)$$

where w_{ij} is the weight of edge $<v_i, v_j>$ and ∞ is a big value that is more than all the weight of the edges and is accepted by the computer. The weight w_{ij} is the measure distance between nodes. If $i=j$, $w_{ij}=0$. The asymmetric communication between nodes can be expressed by the Adjacency Matrix, and then the shortest path between two nodes can be obtained through the Floyd algorithm [20] or the Dijkstra algorithm[21], the complexity of the two algorithms are $O(n^3)$. We denote the shortest path between node i and j as $A\_shortest\_path[i, j]$. Node i and j are connective if $A\_shortest\_path[i, j]$ is less than ∞. It means that there is a path along which the information can be transferred from i to j. Otherwise, node i and j are not connective.

The three situations mentioned before can be distinguished by $A\_shortest\_path[i, j]$ and $A\_shortest\_path[j, i]$:

1) If $A\_shortest\_path[i, j] = A\_shortest\_path[j, i]$ and $A\_shortest\_path[i, j]$, $A\_shortest\_path[j, i]< \infty$, there are two opposite paths through the same nodes.

2) If $A\_shortest\_path[i, j] \neq A\_shortest\_path[j, i]$ and $A\_shortest\_path[i, j]$, $A\_shortest\_path[j, i]< \infty$, there are two opposite paths through different nodes.

3) If $A\_shortest\_path[i, j] >= \infty$ or $A\_shortest\_path[j, i]>= \infty$, there is only one or no path.

D. The Solutions for Localization with Asymmetric Communication

After identifying the communication path between nodes, some solutions can be taken to reduce the influence of the asymmetric communication on node localization. For situation 1), the paths could be kept as the communication between nodes is symmetric. For situation 2), the two nodes are connective and one of them always knows the shortest path from itself to the other. We use the smaller value of $A\_shortest\_path[i, j]$ and $A\_shortest\_path[j, i]$ as the common shortest path for node i and j. For situation 3), the asymmetric communication cannot be avoided as there is at least one node that cannot receive the information from the other one. It is necessary to find that when and how many nodes are in situation 3).

Definition 5: *Strong Connective Graph*. For any two different vertexes v_i and v_j in $V(G)$ of digraph G, if there are always paths from v_i to v_j and from v_j to v_i, G is strong connective graph.

The scenario established in section II.C is used for investigating the ratio of the node pairs in situation 3) to the all nodes with the change of the connectivity of the networks. Denote the ratio of the node pairs in the network that cannot be connected to each other, the node pairs in situation 3) and the all node pairs in the network as r, a and b, respectively. Thus:

$$r = a/b \qquad (3)$$

Fig. 4 indicates the ratio of unconnected node pairs with the change of the average connectivity. The smaller the average connectivity of network, the larger the r is. When the average connectivity is more than 5, the r is lower than 20% and the influence caused by the nodes in situation 3) could be
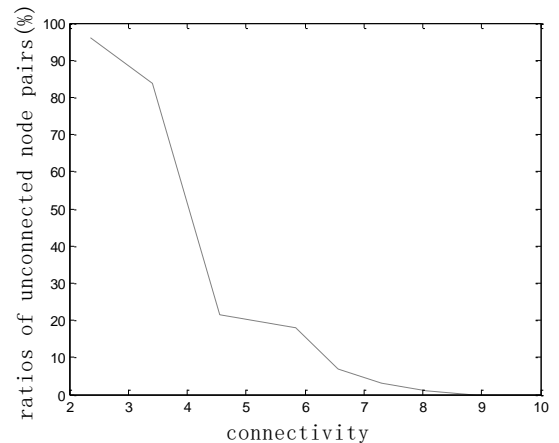
Fig. 4. The ratio of unconnected node pairs VS. the connectivity.
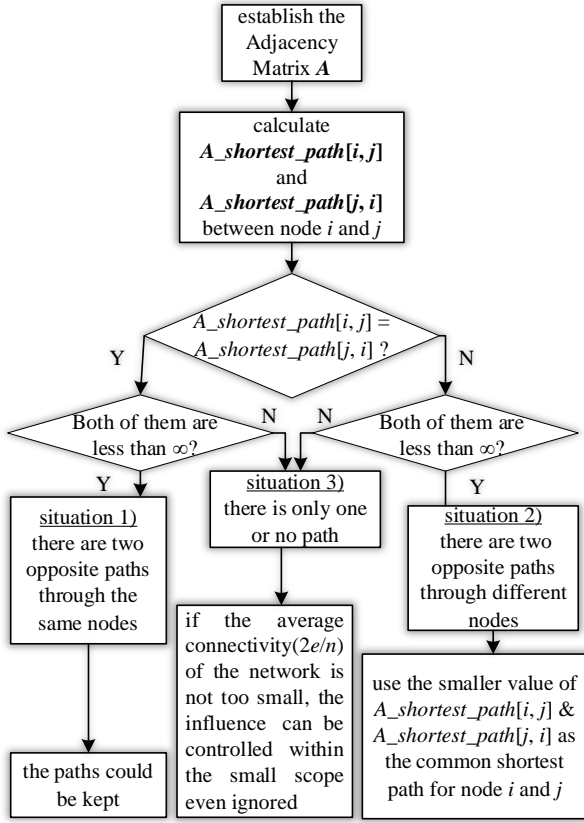
Fig. 5. Flowchart of the developed method.

controlled within the small scope. When average connectivity is more than 9, there is no node in situation 3). It means that the digraph of the senor network is strong connective. Therefore, if the average connectivity of the network is not too small, the influence of the nodes in situation 3) can be controlled within the small scope even ignored. Then it is only needed to consider the nodes in situation 2) for improving the localization performance.

Therefore, the processing flow can be summarized in Fig. 5.

III. SIMULATION EXPERIMENTS

In our simulation experiment, MATLAB is used to simulate WSN and verify the proposed method. Two hundred nodes are randomly deployed in a 300×300 region. The lower bound and the upper bound of the irregular radio is 27 and 53, respectively. The *DOI* is 0.2 and the average connectivity is 8.89. Using this scenario, we perform the node localization 1) in the directed network with asymmetry communication through DV-distance algorithm, 2) in the undirected network with symmetry communication through DV-distance algorithm, 3) in the directed network through the solution mentioned in section II.*D*. In the undirected network, the radio range of the sensor node is set as the average of the upper bound and the lower bound of the radio range in the directed network and the average connectivity is 10.08. The network we establish is strong connective, thus in the directed network, the improvement we need to perform is to chose the letter one between $A\_shortest\_path[i, j]$ and $A\_shortest\_path[j, i]$ as the shortest path for node i and j. This is realizable in both the distributed and centralized localization. The localization performance of the improved solutions is evaluated through comparing the localization errors of the three situations above.
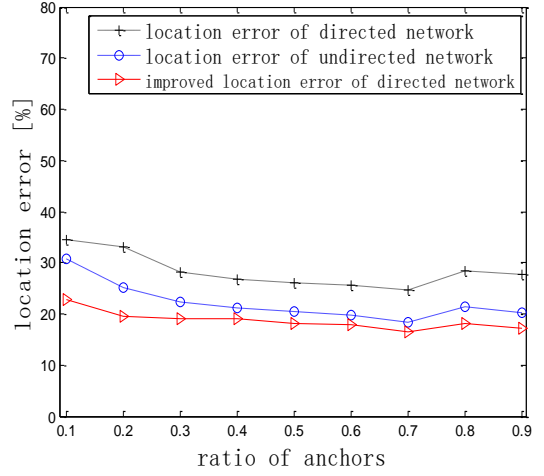
Fig. 6. The comparison of the localization errors for the three situations.

Fig. 6 shows the behavior of the localization error with the change of the anchor proportion. The localization error is obtained by dividing absolute localization error by the radio range. The range of the localization error is 27% ~35% in the directed network, 20%~31% in undirected network, and 17%~23% after improved in the directed network. The localization error decreased by more than 10% through our improved solution. Therefore, the localization algorithm that is suit for undirected network must be improved in the practical application where the radio range is irregular.

The behavior of localization error with the change of *DOI* is shown in Fig. 7(a), the anchor proportion in the simulation is 30%. The upper bound and the lower bound of the radio range is set to be constant when *DOI* changes, but the

(a)

(b)
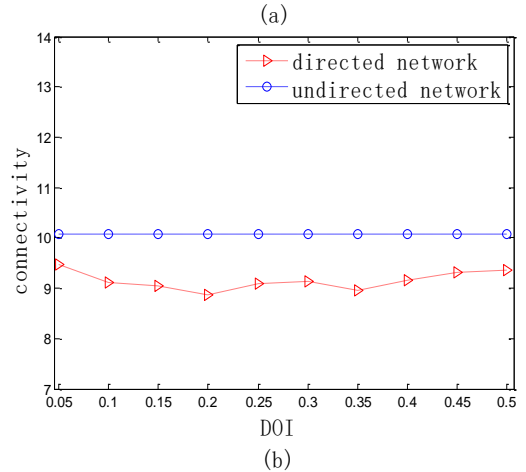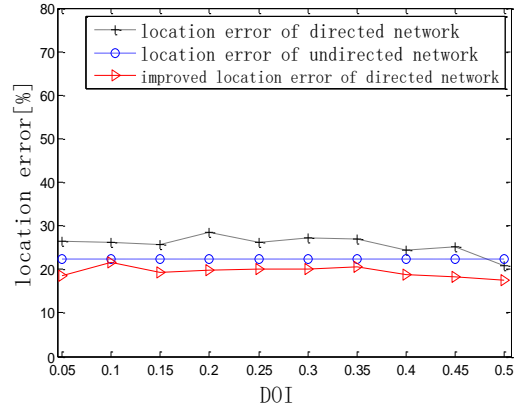
Fig. 7. The localization error VS the DOI.

communication distance varies between the upper bound and the lower bound with the angle. The communication distance changes more greatly if *DOI* is larger. Contrarily, the communication range changes more mildly when *DOI* is
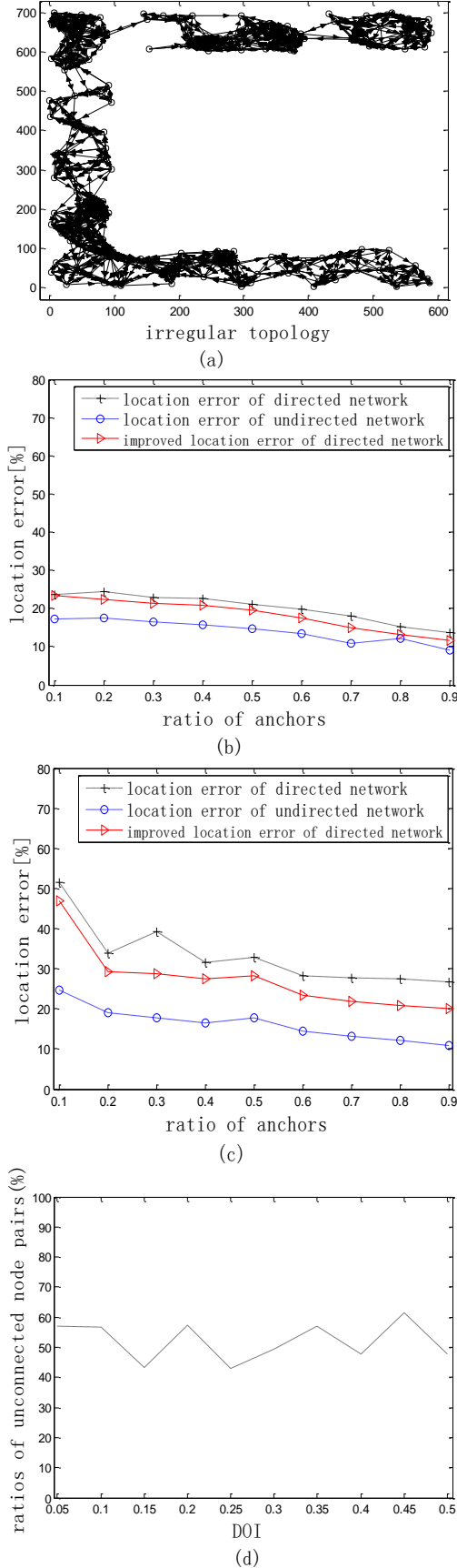
Fig. 8. The localization performance in the 'C' type network.

smaller. It is observed that the change of the localization error is little because the average connectivity is almost invariant when the *DOI* changes (as shown in Fig. 7(b)).

Fig. 8(a) shows the 'C' type network topology. There are 200 nodes and the region of the network is 600×700. Fig. 8(b) shows the behavior of the localization error with the change of the anchor proportion. Here, the network is strong connective and the average connectivity is 12.83 and 14.58 when the network is directed and undirected, respectively. In Fig. 8(c), the network is not strong connective and the average connectivity is 8.46 and 9.28 for the directed network and undirected network, respectively. Fig. 8(d) shows the ratio of the node pairs that cannot contact with each other with the increase of DOI when the network is not strong connective.

The irregular network topology will cause the increase of the node pairs that cannot contact with each other. This also induces the increase of the localization error. In order to reduce the effect of the asymmetric communication and improve the localization accuracy, the average connectivity needs increase. When the average connectivity increases until that the network is strong connective, the localization accuracy will be 11%~24% which is a better scope as shown in Fig. 8(b). When the network is not strong connective, the change scope of the localization error is larger, 19%~52% as shown in Fig. 8(c). The change of DOI has little effect on the node pairs that cannot contact with each other. Therefore, the irregular node distribution is a challenge to the localization when the network is not strong connective.

IV. CONCLUSION

The radio range of the sensor node is not a perfect round. The radio irregularity should be considered for the node localization. In order to perform better in the directed network, the localization algorithm which is suit for the undirected network needs to identify the communication path and modify the shortest path between nodes. The digraph of the node could be used to denote the asymmetric of the information sending and receiving. After be modified, the localization accuracy of the directed network increases more than 10%. The change of DOI has little effect on the localization error of the network. The irregular network topology will induce worse communication between nodes and larger localization error. For example, when the average connectivity is 8.46, the localization error fluctuates between 19% and 52% with the change of the anchor proportion. The strong connective digraph of the irregular topology could be established through increasing the average connectivity.

This paper provides a way to solve the localization problem in the case of asymmetry communication based on Graph Theory. It can quantify the impact of asymmetry, and will reduce the effect of the irregular topology and improve the localization accuracy. Thus, the precision and quality of WSN location applications in aerospace are improved. At present, our work is mainly carried out in MATLAB. In the future, we will build a physical experiment platform to verify the proposed method, and continue to consider the localization problem of isolated nodes at the edge of the WSN.

REFERENCES

[1] SF. Yuan, YQ Ren., L. Qiu, and HF. Mei, "A Multi-Response-Based Wireless Impact Monitoring Network for Aircraft Composite Structures," IEEE transactions on industrial electronics (1982), vol. 63, no. 12, pp. 7712-7722, December 2016.

[2]. F. J. Parrado-García, J. Vales-Alonso, and J. J. Alcaraz, "Optimal Planning of WSN Deployments for In Situ Lunar Surveys," IEEE Transactions on Aerospace and Electronic Systems, vol. 53, no. 4, pp. 1866-1879, August 2017.

[3]. S. Chintalapati, and C.S. Subramanian, "A wireless sensors network system for local multipoint storm surge measurements," Marine Technology Society Journal, vol. 52, no. 4, pp. 32-41, July-August 2018.

[4]. A. A. Ansari, P. Gera, B. Mishra, and D. Mishra, "A secure authentication framework for WSN-based safety monitoring in coal mines. Sadhana (Bangalore)," vol. 45, no. 1, December 2020.

[5]. D. Naranjo-Hernández, J. Reina-Tosina, and L.M. Roa, "Special Issue "Body Sensors Networks for E-Health Applications","" Sensors (Basel, Switzerland), vol. 20, no. 14, pp. 3944, 2020.

[6]. N. Ravi, R. Chitanvis, and M. El-Sharkawy, "Applications of Drones using Wireless Sensor Networks," in Proceedings of the IEEE National Aerospace Electronics Conference, pp. 513-518, 2019.

[7]. T.P. Van, N.P. Van, and T.H. Duyen, "Self-Navigating UAVs for Supervising Moving Objects over Large-Scale Wireless Sensor Networks," International journal of aerospace engineering, no. 2020, pp. 1-20, 2020.

[8]. L. Hsu, C. King, and A. Banerjee, "On Broadcasting in Wireless Sensor Networks with Irregular and Dynamic Radio Coverage," in Proceedings of the International Conference on Parallel Processing, pp. 55-55, 2007.

[9]. S. Phoemphon, C. So-In, and T.G. Nguyen, "An enhanced wireless sensor network localization scheme for radio irregularity models using hybrid fuzzy deep extreme learning machines," Wireless networks, vol. 24, no. 3, pp. 799-819, 2016.

[10]. N. Anand, R. Ranjan and S. Varma, "Radio irregularity model based on received signal strength for three dimensional Wireless Sensor Network," in Proceedings of the 2016 IEEE Region 10 Conference, pp. 2008-2012, 2016.

[11]. M. Singh, and P.M. Khilar, "Mobile beacon based range free localization method for wireless sensor networks," Wireless networks, vol. 23, no. 4, pp. 1285-1300, 2017.

[12]. N. Srikanth, and M. Siva Ganga Prasad, "A Compressive Family Based Efficient Trust Routing Protocol (C-FETRP) for Maximizing the Lifetime of WSN," in Advances in Intelligent Systems and Computing, vol 1049, pp. 69-80, 2020.

[13]. H. Xu, BL. Wang, J. Song, HH. Hong, and XL. Zhang, "An algorithm for calculating coverage rate of WSNs based on geometry decomposition approach," Peer-to-Peer Networking and Applications, vol. 12, no. 3, pp. 568-576, 2019.

[14]. D.S. Sakkari, and T.G Basavaraju, "GCCT: A Graph-Based Coverage and Connectivity Technique for Enhanced Quality of Service in WSN," Wireless Personal Communications, vol. 85, no. 3, pp. 1295-1315, 2015.

[15]. R.R. Swain, T. Dash and P.M. Khilar, "An effective graph-theoretic approach towards simultaneous detection of fault(s) and cut(s) in wireless sensor networks," International Journal of Communication Systems, vol. 30, no. 13, pp. e3273-n/a, 2017.

[16]. B. Sharma, N. Brahma and H. Choudhury, "Graph theory based optimum routing path selection method for wireless sensor network," in proceedings of the International Conference on Advancement of Computer Communication and Electrical Technology, pp. 111-114, 2017.

[17]. K. Thangaramya, R. Logambigai, L. Sairamesh, K. Kulothungan, and A. K. S. Ganapathy, "An Energy Efficient Clustering Approach Using Spectral Graph Theory in Wireless Sensor Networks," in proceedings 2017 2nd International Conference on Recent Trends and Challenges in Computational Models, pp. 126-129, 2017.

[18]. T. He, CD. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-Free Localization Schemes for Large Scale Sensor Networks," in proceedings of the 9th Annual International Conference on Mobile Computing and Networking, pp. 81-95, 2003.

[19]. LJ. Sun, and TF. Chen, "Difference DV_Distance Localization Algorithm Using Correction Coefficients of Unknown Nodes," vol. 18, no. 9, pp. 2860, 2018.

[20]. R.W. Floyd, "Shortest path," Communications of the ACM, pp. 345, 1962.

[21]. E.W. Dijkstra, "A note on two problems in connexion with graphs. Numerische Mathematik," pp. 269-271, 1959.

High Precision Measurement of Dynamic Angular Rate for Turntable in the Calibration Application of Airborne Inertial Sensor

Xuejun He
*Changcheng Institute
of Metrology and Measurement*
Beijing, China
hxj77424@163.com

Aijun Wang
*Changcheng Institute
of Metrology and Measurement*
Beijing, China
waj2044945@163.com

Tianle Zhou
*School of Instrumentation and
Optoelectronic Engineering*
Beihang University
Beijing, China
zpulse@buaa.edu.cn

Xuxing Zhao
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
zhaoxuxing@buaa.edu.cn

Ning Yu*
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
nyu@buaa.edu.cn

Abstract—In the calibration application of airborne inertial sensor, turntable based angular rate measurement is mainly focused on static and constant angular rate test. Relatively little research has been done on the dynamic performance testing under changes in angular rate and angular acceleration rate. A wavelet analysis-based turntable dynamic angular rate measurement method and device are proposed, which can accurately acquire the instantaneous angular rate under turntable speed variation. The dynamic rate experiments were validated for three different angular acceleration situations. The results show that the measurement accuracy of the dynamic angular rate is better than 0.38% during the rapid variation of the velocity, and reaches 0.03% during the steady phase of the rotational speed.

Keywords—dynamic angular rate, wavelet analysis, turntable, circular grating

I. INTRODUCTION

The measurement accuracy of accelerometers, gyroscopes and other airborne inertial sensors directly affect the performance of inertial guidance and navigation systems. The researches are mainly focused on the static measurement and calibration of the inertial sensors, and a more comprehensive testing and calibration techniques has been achieved [1], while relatively few researches are focused on the measurement and calibration of its dynamic performance. In order to evaluate and calibrate the dynamic performance of inertial sensors, high speed and dynamic turntables come into use for providing standard output of angular position, angular rate and angular acceleration. The performance of the turntable directly affects the performance of the calibrated inertial sensor. In the existing turntable performance test and evaluation methods, the static angular position measurement of the turntable is the main method [2-4], and the dynamic performance evaluation of the turntable is mostly under the condition of constant angular rate as the main research point [5]. There are relatively few studies on the changes of the angular rate and angular acceleration rate of the turntable.

The measurement of turntable instantaneous angular velocity (rate) is a key issue in the evaluation of its dynamic performance. At present, the main applied angular measurement methods are laser auto-collimation, ring laser

goniometry and circular grating goniometry. Laser auto-collimation is suitable for static small angle measurement [6], its frequency response is low, the measurement range is small, and the effect is poor for dynamic measurement. The ring laser goniometry has high accuracy, wide dynamic range, and is suitable for dynamic angle measurement [7]. However, the structure of ring laser goniometer is complex and high technological quality is required. At present, only a few institutes such as Xi'an Flight Automatic Control Research Institute, have successfully developed this ring laser goniometer. And limited by the costly price, the ring laser goniometer is relatively less applied in the actual measurement. The circular grating based angular measurement method has higher resolution. It can be applied to the measurement of dynamic angle. Moreover, circular grating is taken as the angle measurement standard in China, and the angular velocity of the turntable measured by the circular grating is convenient for traceability of the value [8]. Therefore, this paper is based on circular grating goniometry to design devices and calculation method for dynamic angular rate measurement.

II. ANGULAR RATE MEASUREMENT PRINCIPLE BASED ON CIRCULAR GRATINGS

In the process of measuring angular rate with circular grating, the circular grating signal can be collected from Mohr
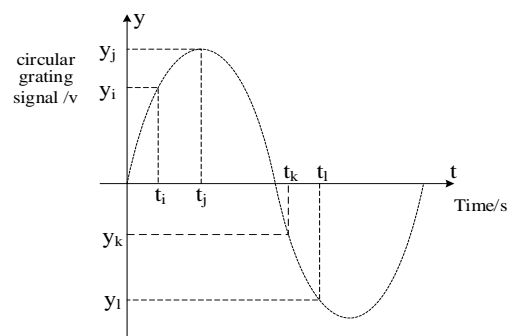
Fig. 1. Diagram of signal waveform produced by circular grating.

stripes generated by the grating through differential and
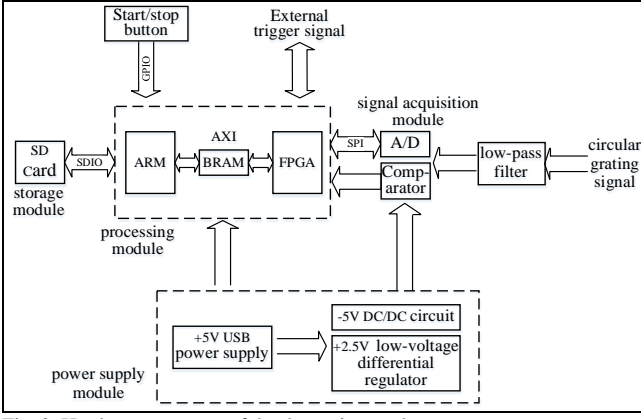
Fig. 2. Hardware structure of the dynamic angular rate measurement system.

Fig. 3. Flowchart of The Dynamic Angular Rate Measurement Method.

amplification circuits, the circular grating signal is a sinusoidal signal, its amplitude remains basically unchanged, the frequency will change with the rotation speed of the turntable [9]. The waveform diagram of the circular grating signal is shown in Fig. 1, and its frequency is directly proportional to the rotational speed (angular rate) of the turntable, that is $f = \dfrac{N_g \cdot \omega}{360^\circ}$. Wherein, f is the circular grating signal frequency, N_g is the line number of circular grating, and ω is the rotation angular rate (unit: $^\circ/s$) of turntable. It can be seen that the turntable rotation angular rate is directly proportional to the output signal frequency of the circular grating. Therefore, by calculating the instantaneous frequency of voltage signal from the circular grating, the dynamic angular rate of rotating turntable will be solved.

III. HARDWARE SCHEME OF DYNAMIC ANGULAR RATE MEASUREMENT SYSTEM

The hardware structure of the dynamic angular rate measurement system is shown in Fig. 2. The hardware of the system consists of four parts, which are signal acquisition module, processing module, storage module and power supply module. The signal acquisition module is responsible for sampling the analog voltage signals output from the circular grating and converting them into digital signals. The processing module is responsible for processing the high speed digital signals and store them to a storage module. The processing module is also interacting with PC system and other external triggers. The storage module is responsible for the backup of the collected digital signals, which can be further used for performing the calculation of dynamic angular rate in PC. The power supply module is responsible for providing power support to the above mentioned system.

The circular grating collects the turntable rotation signal. The signal goes through a low-pass filter to filter out the interference caused by the moving of the turntable motor and the high-frequency noise in the signal acquisition circuit. After that, the output signal of the circular grating is realized through A/D conversion, using a 12-bit analog-to-digital conversion chip in the acquisition module. The data after A/D conversion is processed by ARM and FPGA and stored in a high-speed sd-card. The ARM processor interacts with the PC via the serial port, allowing control of the dynamic angular rate measurement system from the PC. It is also possible to start and stop the dynamic angular rate measurement system using an external trigger signal. At the start of the dynamic angular rate measurement system, an 2 kHz synchronization signal is sent externally to synchronize the system with other instrumentation.
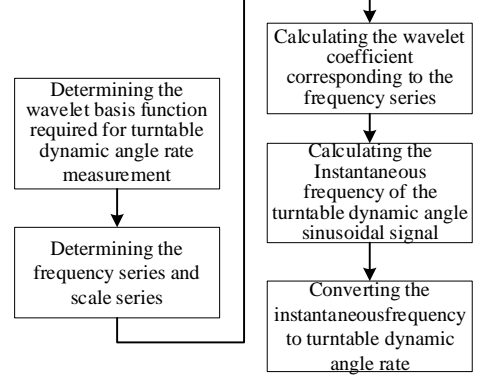
IV. DYNAMIC ANGULAR RATE SOLVING METHOD BASED ON WAVELET ANALYSIS

In order to obtain the dynamic angular rate of turntable, the data collected by the above-mentioned dynamic angular rate measurement hardware system needs to be solved. In this paper, a wavelet analysis-based method for solving the turntable dynamic angular rate is designed. It consists of five steps: determining the wavelet basis function required for turntable dynamic angle rate measurement, determining the frequency series and scale series, calculating the wavelet coefficient corresponding to the frequency series, calculating the instantaneous frequency of the turntable dynamic angle sinusoidal signal, and converting the instantaneous frequency to turntable dynamic angle rate. As shown in Fig. 3. All the steps are performed in PC. The core of the algorithm is to accurately calculate the instantaneous rate [10] of the dynamic angle signal and to establish the conversion model between the instantaneous frequency and the dynamic angle rate.

A. Determining the Wavelet Base Function

Sinusoidal signal from the circular grating is taken as measurement reference to measure the turntable dynamic angular rate. The Morlet function is chosen as the wavelet base function [11] and the function expression is as below.

$$\psi(t) = \exp(-\frac{t^2}{2}) \cdot \exp(j\omega_0 t) \tag{1}$$

Wherein ω_0 is the central frequency of the wavelet base function, t is time, and j is the imaginary unit.

B. Determination of Frequency and Scale Sequences

The measured dynamic angular rate has a frequency range of $[f_{min}, f_{max}]$ ($f_{min}>0$). The frequency range is divided into n-1 equal parts to form the frequency sequence $f_1, f_2, ..., f_{n-1}, f_n$. Wherein $f_1=f_{min}, f_n=f_{max}$. Calculate the scale sequence as $scale_1, scale_2, ..., scale_n$, corresponding to the frequency sequence [12], by using the equation as follows. Wherein F_c is the center frequency of the wavelet base function and F_s is the signal sampling frequency of the A/D module.

$$scale_i = \frac{F_c F_s}{f_i}, \ i = 1, 2, ..., n \tag{2}$$

C. Calculation of Wavelet Coefficients

The sinusoidal signal $y(t)$ outputted by turntable's circular grating is sampled, and a total of p samples are got within a period of time. The samples are represented as $y(t_k)$, $(1 \leq k \leq p, k \in \mathbf{Z})$, wherein $y(t_k)$ is the voltage amplitude of the circular grating signal sampled at t_k moment. Based on the scale sequences $scale_1$, $scale_2$, ..., $scale_n$, respectively, the wavelet basis functions described in step A are transformed by translation and telescoping [13-14] to obtain n scale transformed wavelet basis functions $\psi_1(t), \psi_2(t), ..., \psi_n(t)$, with time domain widths w_1, w_2, ..., w_n. To the circular grating signal $y(t)$ at the moment t_k, its wavelet coefficient a_{ik} for the frequency component f_i is as follow:

$$a_{ik} = \int_{t_k - w_i/2}^{t_k + w_i/2} y(t)\psi_i(t)dt, \ i = 1,2,...,n \qquad (3)$$

The larger the wavelet coefficient a_{ik}, the greater the proportion of the frequency component f_i in $y(t)$ at moment t_k.

Using the matrix $\mathbf{A}_{n \times p}$ to store the wavelet coefficients a_{ik}, $(1 \leq i \leq n, 1 \leq k \leq p, \ i, k \in Z)$, let

$$\mathbf{A}_{n \times p} = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1p} \\ a_{21} & ... & ... & a_{2p} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{np} \end{bmatrix} \qquad (4)$$

Then the wavelet coefficient matrix $\mathbf{A}_{n \times p}$ describes the relative proportion of each frequency component (that is, f_1, f_2, ..., f_{n-1}, f_n) in dynamic angular rate signal.

D. Calculating the Instantaneous Frequency of Sinusoidal Signal of Dynamic Angle

Based on the wavelet coefficient matrix $\mathbf{A}_{n \times p}$, calculate the the instantaneous frequency values f_{1m}, f_{2m}, ..., f_{pm}, corresponding to the moments t_1, t_2, ..., t_p. For the moment t_k, the wavelet coefficients a_{1k}, a_{2k}, ..., a_{nk}, corresponds to the frequency sequence f_1, f_2, ..., f_{n-1}, f_n can be obtained from the matrix $\mathbf{A}_{n \times p}$. The instantaneous frequency f_{km} is calculated by the weighted average method as follows.

$$f_{km} = \frac{\sum_{i=1}^{n} f_i a_{ik}}{\sum_{i=1}^{n} a_{ik}} \qquad (5)$$

Thus, the relation between the measurement time and instantaneous frequency (t_k, f_{km}) is established.

E. Converting Instantaneous Frequency to Turntable Dynamic Angular Rate

Assume the total lines of turntable circular grating as N. Then a 360° rotation of the turntable will output N sinusoidal signals. Convert the instantaneous frequency of the sinusoidal signal to the dynamic angular rate of the turntable, and the conversion relationship is as follow.
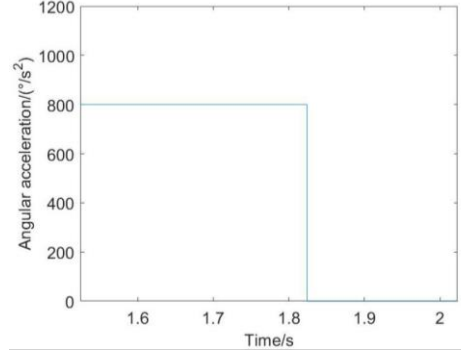
Fig. 4. Variation curve of angular acceleration (constant angular acceleration).
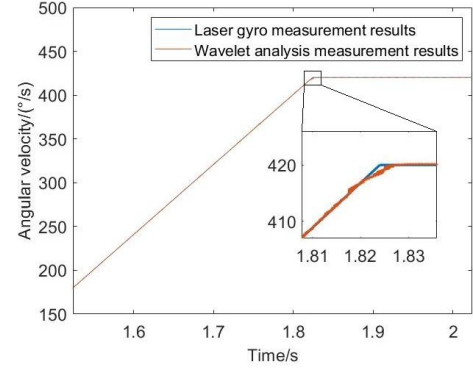
Fig. 5. Angular rate measurement results (ZLG vs Proposed method).
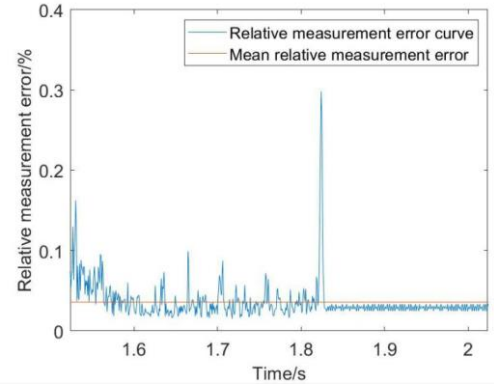
Fig. 6. Relative measurement error of angular rate.

$$\upsilon_k = \frac{f_{km}}{N} \times 360° \qquad (6)$$

This finally realizes the measurement of turntable dynamic angular rate, which varies dynamically with time.

V. DYNAMIC ANGULAR RATE MEASUREMENT EXPERIMENT

In order to verify the proposed dynamic angular rate measurement method, a variable rate dynamic test of turntable was conducted. A model of high dynamic turntable from the Changcheng Institute of Metrology and Measurement is used. The sampling frequency of analog-to-digital conversion module is 250kHz, which is much higher than the output frequency of circular grating and satisfies the changes in rotation speed or frequency without affecting the measurement outcome. The processor zynq-7020 is used as a digital signal processing module. The zero-lock-area quad-frequency ring laser goniometer (ZLG) developed by the Xi'an

flight automation control institute to carry out synchronous measurement of turntable with variable rate. The angular rate measurement range of this ZLG is higher than 450°/s, the angular acceleration rate measurement range is higher than 2000°/s², the angular resolution is 0.01 arcsec (1σ), the scale factor nonlinearity is better than 0.1×(1σ), and the dynamic angular measurement accuracy is higher, which can be used as a standard to compare with the calculation results of the proposed method in our paper .

In the test, the turntable was started from zero speed, and the angular acceleration was in the form of constant, constant increase and sinusoidal decrease, which are the three main types of variable speed used in the mainstream motor operation currently, so as to verify the measurement effect of the proposed method when the rotational speed of the turntable changes. When the angular rate finally reaches 420°/s, the turntable is kept running at a constant speed rate and the rotational speed does not change.

A. Constant Angular Acceleration

The turntable starts its motion from zero velocity and maintains a constant angular acceleration of 800°/s², and the angular rate gradually increases. The angular acceleration curve, the angular rate measurement curve based on wavelet analysis method, and the relative measurement error curve are shown in Fig. 4, 5, and 6, respectively. It shows the dynamic angular rate measurement method based on wavelet analysis can accurately track the changes in the dynamic angular rate. At the stage of angle rate increment, the relative measurement error of the dynamic angular rate based on the wavelet analysis method is kept below 0.15%, which can achieve a high dynamic measurement accuracy. At the moment t=1.82s, the turntable rate enters the stable stage, no longer accelerating but maintaining a uniform motion. At this moment, the measurement error is slightly larger, reaching a maximum of 0.3%, which is closely related to the turntable's motion state mutation. Thereafter, the turntable rotates at a constant speed, and the relative measurement error is reduced to about 0.03%. The average relative measurement error during the whole movement is about 0.034%.

B. Constant Increase in Angular Acceleration

The turntable starts its motion from zero velocity with an initial value of angular acceleration of 600°/s², and the angular acceleration gradually increases at a rate of 800°/s³. The angular acceleration variation curve, the angular rate measurement curve based on the wavelet analysis method, and the relative measurement error curve are shown in Fig. 7, 8, and 9, respectively. Similar to the former case, the relative measurement error of the dynamic angular rate based on wavelet analysis method remains below 0.3% (mainly distributed in 0.03% ~ 0.15%) at the stage of angular rate increment. At the moment when the angular rate begins to enter the stable phase, the angular acceleration increases to more than 1000°/s², and the relative measurement error reaches a maximum value of 0.38% because the abrupt change of the turntable's motion state. Thereafter the turntable rotates at a constant speed, and the relative measurement error is reduced to about 0.03%. The average relative measurement error of the whole process is slightly higher than that of the constant angular acceleration, about 0.042%.

C. Sinusoidal Decrease in Angular Acceleration

The angular acceleration decreases in a sinusoidal form, with an initial value of 1200°/s² and a final value of 0°/s². The
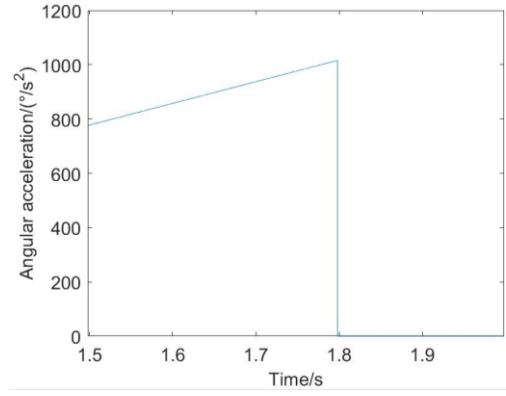
Fig. 7. Variation curve of angular acceleration (constant increase in angular acceleration).
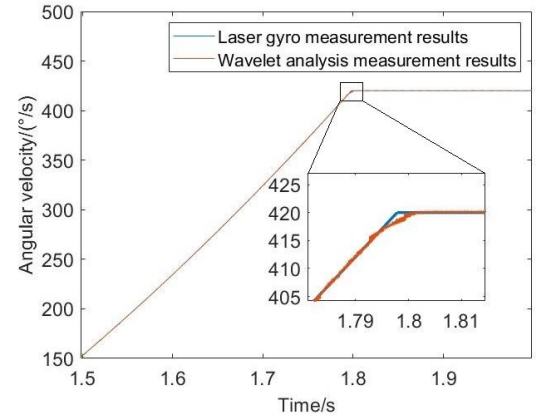
Fig. 8. Angular rate measurement results (ZLG vs Proposed method).
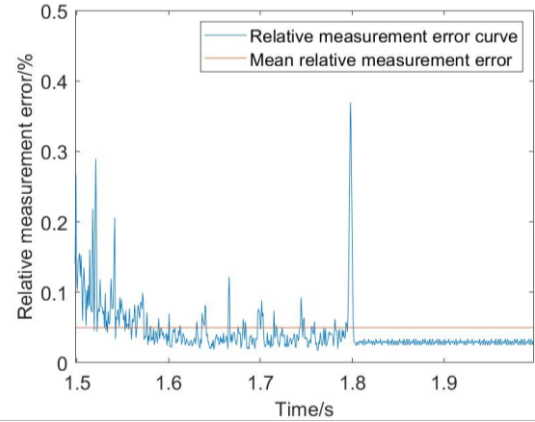
Fig. 9. Relative measurement error of angular rate.

angular acceleration variation curve, the angular rate measurement curve based on wavelet analysis, and the relative measurement error curve are shown in Fig. 10, 11, and 12, respectively. Angular acceleration and angular rate are both changing like the sinusoidal curve. In the angular rate incremental stage, the relative measurement error of the dynamic angular rate based on wavelet analysis method is kept at 0.01 %~ 0.12%. At the moment when the angular rate begins to enter the stable phase, the relative measurement error is also less than 0.1%, better than the previous two cases. It is because the angular acceleration keeps going down and decreases to zero at the moment of entering the steady state. The average angular acceleration is also lower than case B. Thereafter, when the turntable rotates at a constant speed, the relative measurement error is reduced to about 0.03%. The average relative measurement error during the entire motion is about 0.032%.
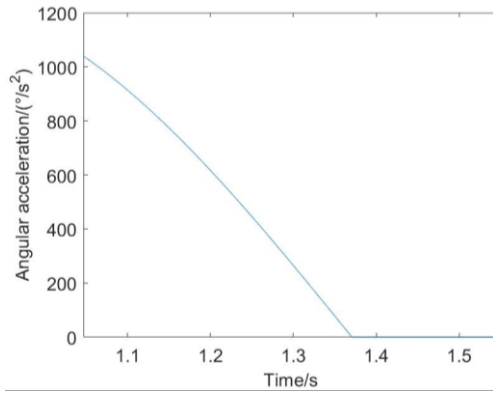
Fig. 10. Variation curve of angular acceleration (sinusoidal decrease in angular acceleration).
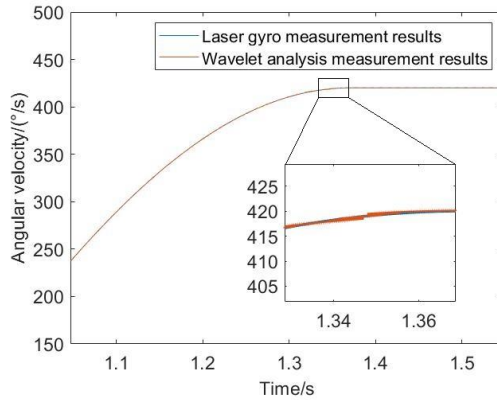
Fig. 11. Angular rate measurement results (ZLG vs Proposed method).
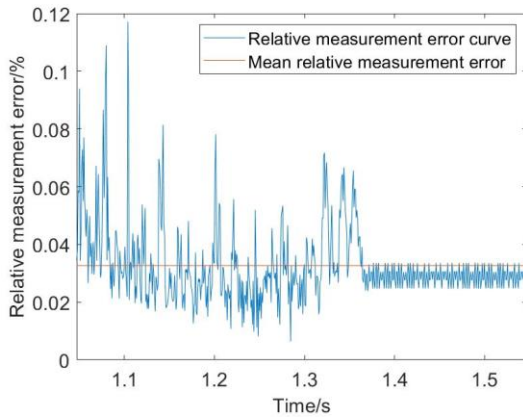
Fig. 12. Relative measurement error of angular rate.

VI. CONCLUSION

This paper designs a high precision measurement system for dynamic angle rate, designs the hardware scheme of measurement system, proposes a wavelet analysis-based dynamic angle rate measurement method, and finally realizes the accurate measurement of dynamic angle rate. Experimental results show that the measurement device and measurement method can accurately measure the instantaneous angular rate during turntable rotation, and the average relative measurement error is less than 0.042%, which makes up for the shortage of traditional static angular rate measurement and has a high application value in precision turntable dynamic velocity measurement for the calibration of airborne inertial sensor.

REFERENCES

[1] QE. He, CJ. Zeng, and X. He, "Calibrating accelerometers for space-stable inertial navigation systems at system level," Measurement, vol. 127, no. 1, pp. 472-480, 2018.

[2] RS. Zhu, XL. Cao, and ZH. Zhong, "A probe into test method of turntable angle position accuracy," Navigation and Control, vol. 12, no. 2, pp. 73-77, 2013.

[3] YY. Cai, J. Yang, and JK. Li, "The Error Model and Experiment of Measuring Angular Position Error Based on Laser Collimation," International conference on optical instruments and technology, vol. 10621, Article number: 1062114, 2017.

[4] HH. Yang, WL. Feng, F. Liu, J. Zhou, and Z. Ma, "Error analysis and compensation of angular position measuring device for turntable working surface," Chinese Journal of Scientific Instrument, vol. 38, no. 5, pp. 1184-1189, 2017.

[5] SQ. Tang, JS. Liu, and XL. Zhang, "Research on dynamic goniometry of high-accuracy motion simulator," Journal of Astronautic Metrology and Measurement, vol. 31, no. 2, pp. 8-12, 2011.

[6] YL. Chen, S. Yuki, T. Jun, N. Kazuki, and M. Hiraku, "Laser autocollimation based on an optical frequency comb for absolute angular position measurement," Precision Engineering, vol. 54, no. 1, pp. 284-293, 2018.

[7] JX. Wang, LF. Chen, and X. Fu, "A high precision dynamic angle measuring device and method", China: CN105091844A, 2015-11-25.

[8] G. Chen, "Improving the angle measurement accuracy of circular grating," Review of Scientific Instruments, vol. 91, no. 6, Article number: 065108, 2020.

[9] Y. Su, Q. Wang, F. Yan, X. Liu, and YM. Huang, "Subdivision Error Analysis and Compensation for Photoelectric Angle Encoder in a Telescope Control System," Mathematical Problems in Engineering, vol. 2015, no. 16, pp. 1-9, 2015.

[10] P. Chen, QM. Li, and T. Zhao, "Advances and Trends in Instantaneous Frequency Estimation Methodology," Electrical Measurement & Instrumentation, vol. 07, pp. 1-7, 2006.

[11] H. Yi, P. Ouyang, and T. Yu, "An algorithm for Morlet wavelet transform based on generalized discrete Fourier transform", international journal of wavelets multiresolution and information processing, vol. 17, no. 5, Article number: 1950030, 2019.

[12] ZK. Peng, YY. He, and FL. Chu, "Using Wavelet Scalogram for Vibration Signals Analysis," vol. 38, no. 3, pp. 122-126, 2002.

[13] HQ. Wang, CY. Shang, RP. Gao, and ZN. Li, "An Improvement of Wavelet Shrinkage Denoising via Wavelet Coefficient Transformation," Journal of Vibration and Shock, vol. 30, no. 10, pp. 165-168, 2011.

[14] J. Zhou, YP. Chen, C. Zhou, and J. Liang, "The Application of De-noise Based on Compromise Algorithm Between Soft-thresholding and Hard-thresholding of Wavelet Coefficient for Fault Location," Automation of Electric Power Systems, vol. 35, no. 01, pp. 65-68, 2005.

Design Method of the Heating System for Controllable Heating Humidity Sensor

Chun Hu
School of Electronics and Information
Engineering
Beihang University
Beijing, China
buaa_hc@buaa.edu.cn

Dapeng Li
School of Electronics and Information
Engineering
Beihang University
Beijing, China
dapengLI@buaa.edu.cn

Peng Peng
Service Robot Testing Department
Chongqing Dexin Robot Testing
Center
Chongqing, China
272508585@qq.com

Dezhi Zheng*
Innovation Institute of Frontier Science
and Technology, School of Instrument
Science and Opto-electronics
Engineering
Beihang University
Beijing, China
zhengdezhi@buaa.edu.cn

Abstract—**Humidity sensors are prone to condensation in high humidity and low temperature environments, reducing the sensor's measurement performance. Aiming at the current humidity sensor heating system that may not reach the target temperature due to open-loop control and high power consumption, a closed-loop control heating system for the humidity sensor is designed to perform real-time heating and temperature detection of the humidity-sensitive structure. The finite element analysis method is used to conduct thermal analysis on the heating structure, which obtains the temperature distribution of the sensor heating area, guides the heater structure design and gains the location distribution of temperature sensors. A variable parameter fuzzy PID (Proportion Integration Differentiation) control method with variable target temperature is proposed to perform closed-loop control on the temperature of the sensor's sensitive structure. The heating target temperature is set as a function of the ambient temperature. Also the coefficient of PID changes with error. Thus, the power consumption of the system has been effectively reduced and the heating speed is apparently increased. Simulation experiments show that the heating system designed for controllable heating humidity sensor has uniform temperature distribution. When the ambient temperature is $-70℃$, the difference between upper and lower temperatures is only $0.0878℃$, with a low power consumption($5 \times 10^6 W/m^3$). At the same time, the overregulation of the temperature control system is low, which is 7.227%. The system has high environmental adaptability and temperature stability.**

Keywords—humidity sensor, heating system, finite element analysis, fuzzy PID

I. INTRODUCTION

In the fields of weather detection, agricultural production, industrial control, medical monitoring, food storage, etc., humidity detection is actively demanded. In daily production and life, humidity measurement plays an increasingly important role in daily production and life, or even an indispensable character. However, when the measured ambient temperature is higher than the sensor temperature, especially in harsh environments such as places with high altitude and low temperature, condensation often occurs on sensor's surface, which greatly reduces the measurement accuracy of the sensor, affects the response speed, and even seriously invalidate the sensor. Facing the above problems,

the heating humidity sensor is currently the focus of research by scholars [1-6]. Although they solve the condensation problem to varying degrees, the designed heating system uses an open-loop control method. These methods either fit the heating power and the ambient temperature into a linear relationship, or apply a fixed power with fixed period to the heater, without detecting the heating target temperature, which results in poor adaptability. Therefore these methods may not only cause the target temperature unreached, but also result in high power consumption.

In view of the above situation, a design method of heating system for a controllable heating humidity sensor is proposed, which can conduct real-time heating and temperature detection of humidity-sensitive structures through closed-loop control method. In order to better guide the heater structure design and study temperature sensor location distribution, the heating structure is thermally analyzed by the finite element analysis method, and a more reasonable heater structure is designed. At the same time, a variable parameter fuzzy PID (Proportion Integration Differentiation) closed-loop control method with variable target temperature is proposed to perform closed-loop control on the temperature of the sensor's sensitive structure. The heating target temperature is set as a function of the ambient temperature, thereby effectively reducing the power consumption of the system and improving the temperature control accuracy and heating speed. The simulation result shows that the designed heating system of controllable heating humidity sensor has uniform temperature distribution, low power consumption and high temperature stability.

II. STRUCTURE DESIGN AND SIMULATION ANALYSIS OF HUMIDITY SENSOR HEATER

Since the moisture that the air can hold is proportional to the temperature, when the moist air encounters a lower temperature surface, the temperature of the moist air will also drop accordingly, so its ability to hold moisture is also reduced. If the temperature is lower than this limit, moisture will precipitate on the surface, causing condensation [7]. Based on the above considerations, the heater structure of the humidity sensor is designed. By heating the sensor to make the surface temperature higher than the ambient temperature, the measurement accuracy of the sensor is improved.

*Dezhi Zheng is the corresponding author. (e-mail: zhengdezhi@buaa.edu.cn).

A. Theoretical Analysis of Humidity Sensor Heat Transfer

For general humidity sensor with heater, the sandwich structure design is adopted [8]. The main body is composed of humidity sensor, heater, and silicon base layer from top to bottom, which are closely attached to each other in sequence, as shown in Fig. 1.
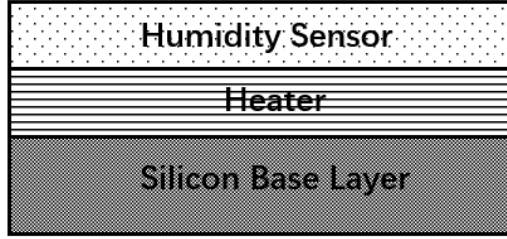
Fig. 1. Schematic diagram of the humidity sensor heating structure.

The heater is located between the sensor structure and the silicon base layer, so that the humidity sensor is evenly heated. The silicon base layer is located in the lowest layer, and its thermal conductivity is low, which can play a certain thermal insulation effect and ensure the heating efficiency of the heater.

In order to better analyze the performance of the humidity sensor with heater, a thermal analysis model of the humidity sensor is constructed, and the structure of the humidity sensor is abstracted as a rectangular parallelepiped, where the heating area is S and the model height is H, as shown in Fig. 2.
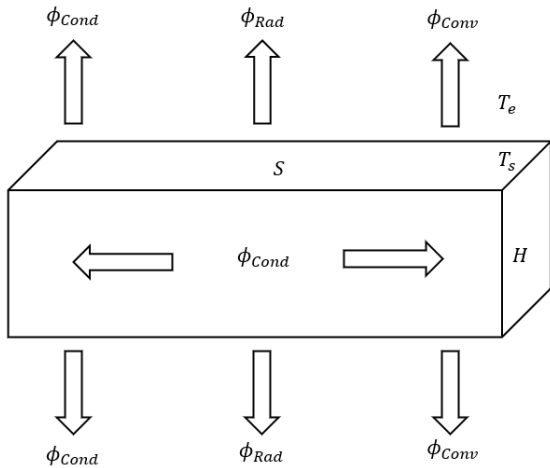
Fig. 2. Theoretical model of the humidity sensor heating structure.

For humidity sensors, there are three heat transfer methods: heat conduction ϕ_{Cond}, heat radiation ϕ_{Rad}, and heat convection ϕ_{Conv} [9]. The theoretical model of heat transfer is shown in Figure 2 above.

Among them, T_s is the surface temperature of the sensor, and T_e is the ambient temperature. The total heat dissipation of the entire sensor structure can be expressed by (1):

$$\phi = \phi_{Cond} + \phi_{Rad} + \phi_{Conv} \quad (1)$$

Heat conduction is the process of transferring heat energy from high temperature to low temperature. The heat conduction equation of humidity sensor can be expressed by (2):

$$d\phi_{Cond} = -\lambda S(dT/dx) \quad (2)$$

In the formula, λ is the thermal conductivity, S is the area of the heating zone, and dT/dx is the temperature gradient.

For this model, the heat transfer can be obtained by integrating both ends of equation (2) with x as shown in (3):

$$\phi_{Cond} = \lambda S(T_s - T_e)/H \quad (3)$$

Thermal radiation refers to the phenomenon that an object radiates electromagnetic waves due to its temperature. The thermal radiation equation of the humidity sensor can be expressed by (4):

$$\phi_{Rad} = 2\varepsilon\sigma S(T_s^4 - T_e^4) \quad (4)$$

In the formula, ε is the emissivity of the object, and it is assumed that the sensor is an ideal blackbody $\varepsilon = 1$ [10], and σ is Boltzmann's constant $\sigma = 5.67 \times 10^{-8} W/(m^2 \cdot K^4)$ [11].

Heat convection is the process of heat transfer through the flowing medium. The heat convection equation of the humidity sensor can be expressed by (5):

$$\phi_{Conv} = \mu(T_s - T_e) \quad (5)$$

In the formula, μ is the air convection coefficient, generally $\mu = 10 W/(m^2 \cdot K)$ [12].

In this system, in order to improve the heating efficiency and maximize the heat generation rate of the humidity sensor surface temperature, single crystal silicon with a larger heat transfer coefficient is selected as the substrate layer. Assuming system $S = 28 \times 10^{-6} m^2$, $H = 0.67 \times 10^{-3} m$, ambient temperature $T_e = -65°C$, humidity sensor surface temperature $T_s = -70°C$, single crystal silicon heat transfer coefficient $\lambda = 157 W/m \cdot °C$, at this time, the thermal conductivity of the upper surface of the system is:

$$\phi_{Cond} = \lambda S(T_s - T_e)/H = 32.81 W/m^2 \quad (6)$$

Similarly, the heat conductivity of the remaining two surfaces can be calculated to be $0.30 W/m^2$ and $0.92 W/m^2$.

Since the thermal radiation of the system is related to Boltzmann's constant, the thermal radiation of the system is too small to be ignored in the calculation of the total heat dissipation of the system.

According to the heat convection equation, the heat convection flow of the system is:

$$\phi_{Conv} = \mu(T_s - T_e) = 50 W/m^2 \quad (7)$$

From the above calculation, it can be seen that the heat convection is the maximum heat loss method of the system heat. According to this method, the heat conduction of all surfaces of the system is calculated separately and added to the heat convection of the system, and the total heat dissipation of the system is about:

$$\phi = \phi_{Cond} + \phi_{Conv} = 118.06 W/m^2 \quad (8)$$

B. Design and Analysis of Humidity Sensor Heater

After the theoretical model of the sensor heat transfer is obtained, the structure of the heater needs to be designed so that the humidity sensor electrode is heated evenly and the heated area is wide to achieve the best heating effect. In order to explore a more reasonable heater structure, several conventional heater structures are designed, and the specific dimensions of the structure are marked, as shown in Fig. 3.

Use Ansys finite element analysis software to conduct thermal analysis on the above structure, and mesh the structure through auto mode. Assume that the ambient temperature is

$-70℃$, the substrate material is silicon, and the heater material is platinum. The theoretical analysis conclusion is used to guide parameter setting, in which the proposed convection coefficient is $10W/(m^2 \cdot K)$, and the heat generation rate of the heater is $5 \times 10^6 W/m^3$. In order to simplify the simulation model, it is assumed that the coefficient changes of the thermal conductivity of sensor material is ignorable, and steady-state thermal analysis is performed on the model. Fig. 4 shows the results of the finite element analysis of each structure.

a) The first heater structure

b) The second heater structure

c) The third heater structure

d) The fourth heater structure

Fig. 3.   Four heater structures.

a) Thermal analysis of the first heater structure

b) Thermal analysis of the second heater structure

c) Thermal analysis of the third heater structure

d) Thermal analysis of the fourth heater structure
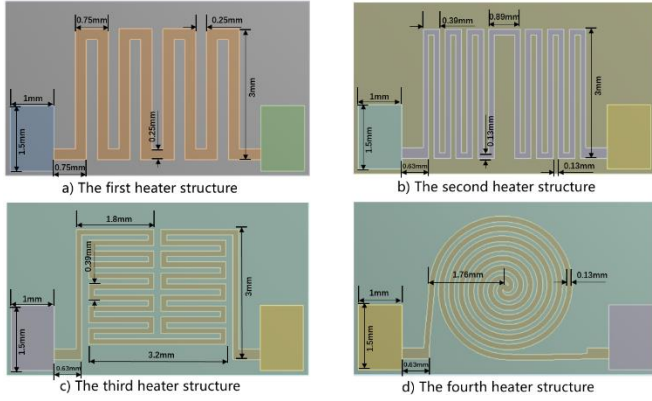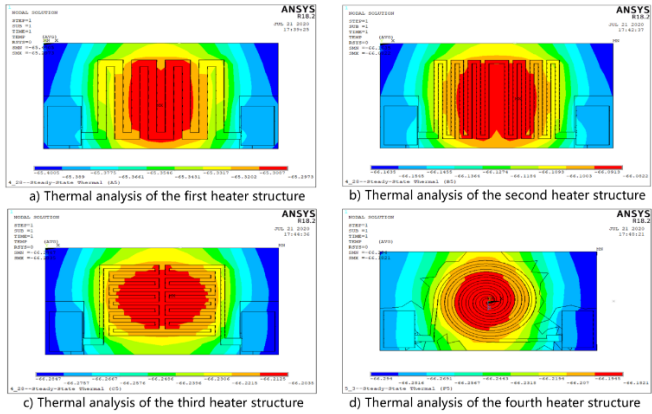
Fig. 4.   Thermal analysis of four heater structures.

Through the thermal analysis results of the sensor surface, it can be seen that the temperature of the above four structures after heating can reach about $-65℃$, which is higher than the ambient temperature set as $-70℃$. It not only meets the requirement of eliminating sensor condensation, but also consumes only $5 \times 10^6 W/m^3$ as heating power, which is much lower than $1.16 \times 10^{11} W/m^3$ in the previous work [2]. In order to further clarify the temperature distribution of each structure, the centerline of the above four structures along the length direction is taken as the path, and the temperature distribution curve under the path is drawn as Fig. 5.

It can be seen that the first three structures have the problem that the highest temperature is not in the center position, which will cause uneven heating of the humidity sensor electrode and affect the accuracy of the measurement. On the basis of the above structure, optimization and improvement are carried out, and the heater structure shown in Fig. 6 is designed.

a) Temperature distribution curve of the first structure

b) Temperature distribution curve of the second structure

c) Temperature distribution curve of the first structure

d) Temperature distribution curve of the forth structure
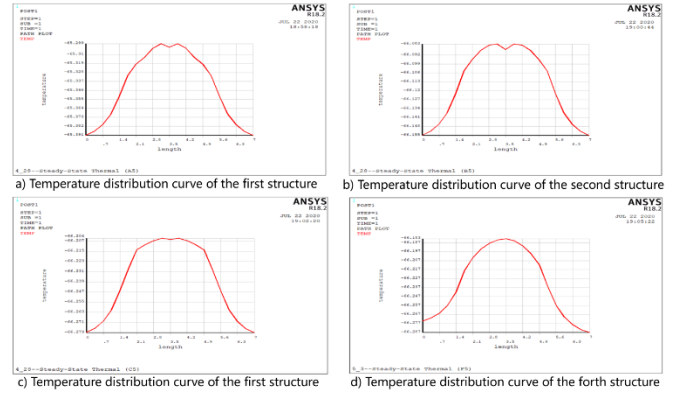
Fig. 5.   Temperature distribution curves of four heater structures.
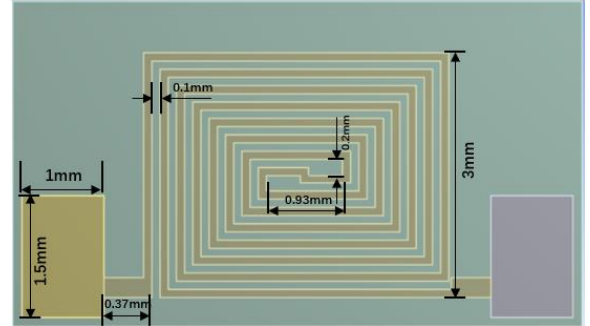
Fig. 6.   The fifth heater structure.

The heater structure size is shown in the figure, the total volume $V = 1.0254mm^3$, the upper surface heating area $S = 6.1527mm^2$, the finite element analysis of the structure shows the surface heat distribution as shown in Fig. 7.
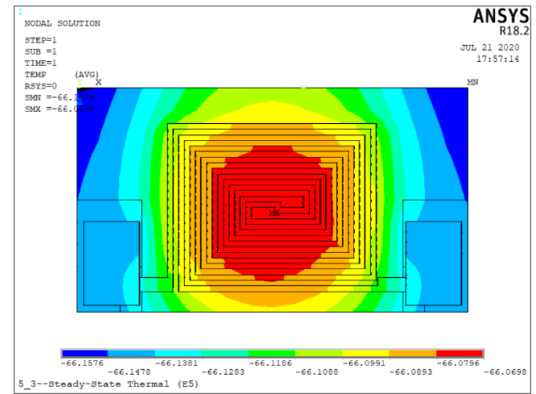
Fig. 7.   Thermal analysis of the fifth heater structure.

The simulation results show that the designed heater structure can uniformly heat the sensor surface. When the ambient temperature is still $-70℃$, the average temperature of the sensor surface is heated to $-66.1137℃$, and the difference between the highest temperature and the lowest temperature of the sensor's surfaces is only $0.0878℃$, which at the same time, it can meet the needs of normal work under the condition of lower power. In order to better explain the advantages of the designed heater structure, the horizontal and vertical temperature distribution curve of the structure is made as Fig. 8.

a) Lateral temperature distribution curve   b) Longitudinal temperature distribution curve
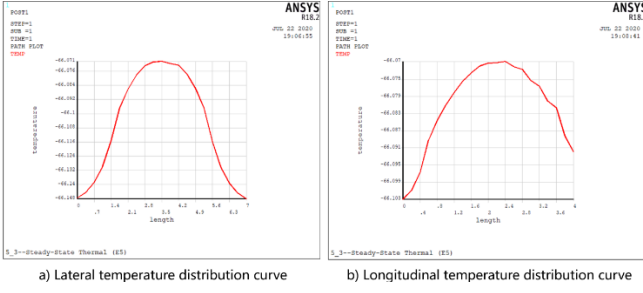
Fig. 8. The fifth structure's horizontal and vertical temperature distribution curve.

As shown in the figure, the fifth heater structure design ensures that the core temperature is the highest temperature, while the temperature distribution is more uniform. Further, the designed fifth structure and the first four structures are compared from different angles, as shown in Tab. 1.

TABLE I.   DIFFERENT HEATER STRUCTURE PERFORMANCE COMPARISON TABLE

Structure Name	Volume /mm³	Upper Surface Area /mm²	Highest Temperature /°C	Lowest Temperature /°C	Temperature Difference /°C	Cover Length /mm
1	1.1154	6.6926	-65.297	-65.401	0.104	4.0
2	0.9703	5.8126	-66.082	-66.164	0.082	4.6
3	0.9288	5.5730	-66.204	-66.285	0.081	4.4
4	0.9357	5.6003	-66.182	-66.294	0.112	3.7
5	1.0254	6.1527	-66.070	-66.158	0.088	4.7

Among the table, the ambient temperature set by the system is $-70°C$. Compared with the fourth disc structure, the second and third serpentine structures have a more even temperature distribution and a smaller temperature difference. When observing the temperature distribution curve of each structure, it can be found that the disc structure can ensure that the highest temperature point of the structure is at the center. At the same time, by comparing the two structures, it can be seen that the finer structure distribution saves material while ensuring better heating effect and wider coverage. The coverage here refers to the length distance contained in the temperature distribution curve when the maximum temperature drops by $0.05°C$. The fifth structure is designed after fully absorbing the advantages and making up for thr disadvantages of the first four structures. While ensuring the highest temperature in the central area, the structure makes the sensor evenly heated and the heating coverage becomes wider. Therefore, it is more reasonable to select the fifth structure as the heater structure of the humidity sensor with heater.

C. Temperature Sensor Heater Distribution

Using the designed structure as the humidity sensor heater structure, in the process of thermal analysis, the highest and lowest temperature distribution positions of the entire structure are obtained, which can be used to guide the position distribution of temperature sensors used in the system. The schematic diagram of the position distribution of the temperature sensor of the sensitive structure is shown in Fig. 9. By observing the temperature distribution curve of the structure, the humidity sensor sensitive structure is selected to be placed in the highest temperature central area of the structure, and the lower part is selected as the distribution position of the temperature sensor T_s of the sensitive structure to ensure that the measured temperature is close to the real temperature of the sensitive structure to the greatest extent. In

order to prevent the influence of heating on the ambient temperature sensor, the ambient temperature sensor cannot be placed in the heating area, and it must be insulated from the heating area. Through this position distribution, the measurement error caused by the temperature of the heater itself is avoided, and a reasonable guarantee is provided for the follow-up work.
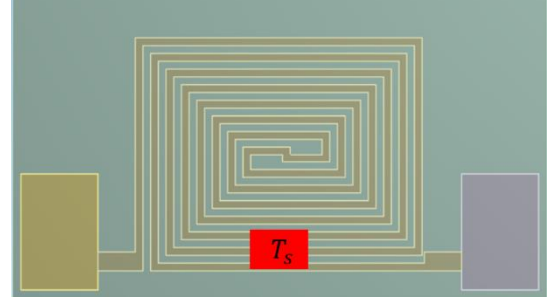
Fig. 9. Schematic diagram of system temperature sensor location distribution.

III. VARIABLE PARAMETER SELF-TUNING FUZZY PID CONTROL METHOD

In the control method of the heater, the open loop control is mostly used in the past work, the target temperature is set at a constant temperature, and the heating power is adjusted to reach the target temperature. However, this method consumes a lot of power, and may not reach the preset target temperature due to the influence of the external environment, causing humidity measurement errors. This paper proposes a variable parameter self-tuning fuzzy PID closed-loop control method with variable target temperature. Through the setting of fuzzy rules, the parameters of the system are continuously adjusted and the adjustment speed is accelerated. In order to reduce the power consumption of the system, the target temperature is only set $5 - 10°C$ higher than the ambient temperature. The ambient temperature and the sensitive structure temperature are measured and fed by two set temperature sensors, and the target temperature is adjusted by self-adjusting in time, which reduces the error and the heating power consumption at the same time.

A. Principle of Fuzzy PID Control Method

PID control method has strong robustness and is widely used in industrial control [13]. However, in this system, the target temperature may be changing at any time, the parameters of traditional PID control method are fixed, and it is easy to cause large overshoot and long response time when used in heaters with lag link. The fuzzy PID control method can detect and analyze uncertain conditions, parameters, delays, interference and other factors in the control process [14]. Using this method not only maintains the robustness of the traditional PID method, but also brings advantages such as good flexibility and high control accuracy. According to the designed fuzzy PID control strategy, the structure of the fuzzy PID controller shown in Fig. 10 is designed.
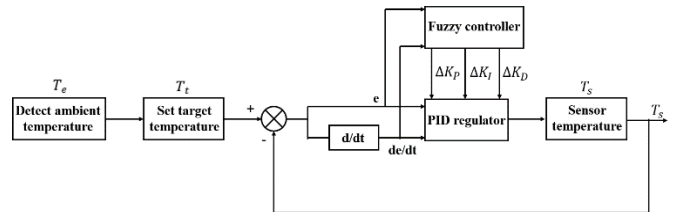
Fig. 10. Fuzzy PID controller structure.

The system continuously detects the ambient temperature through the installed ambient temperature sensor and sets it as the target temperature to achieve the effect of variable target temperature. The input of the fuzzy controller consists of two parts: one is the deviation e between the sensor detection temperature T_s and the preset target temperature T_t, the other one is the time differential de/dt of the deviation. The ΔK_P, ΔK_I, ΔK_D are input into the PID regulator to adjust the heating power, and then control the heater of the controlled object. The temperature is continuously detected by the sensor temperature T_s to determine the deviation from the target temperature, and then the parameters are continuously adjusted until the temperature reaches the preset temperature.

B. Fuzzy PID Control Simulation and Result Analysis

Fuzzy controller is the key to fuzzy PID control. The input variables e and de/dt of the system will be represented by fuzzy subsets [15], and its domain is set to [-6 6]. Express the language values of the output variables ΔK_P, ΔK_I, and ΔK_D of the system as {(NB), (NM), (NS), (ZO), (PS), (PM), (PB)}, a total of seven level [16], and set the corresponding domains to [-6 6]. The input and output are compiled with certain fuzzy rules and imported into the fuzzy controller [17]. The simulation structure diagrams of fuzzy PID and traditional PID are established in the SIMULINK module, as shown in Fig. 11 and Fig. 12.
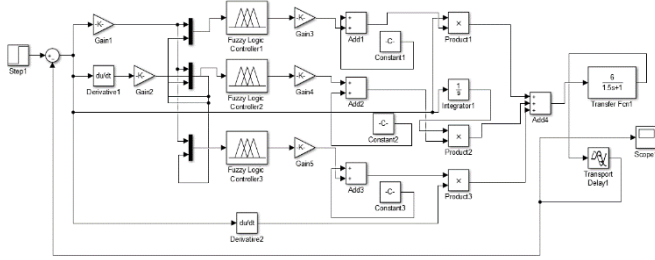
Fig. 11. Fuzzy PID control system simulation structure diagram.
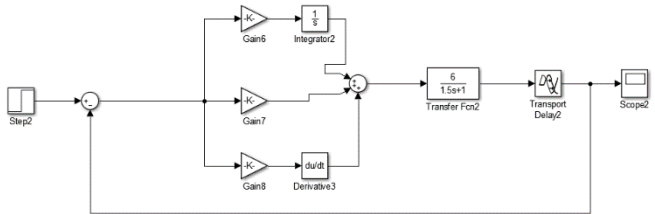
Fig. 12. Simulation structure diagram of traditional PID control system.

For the three gain parameters of PID, the parameters of fuzzy PID are calculated as follows [18]:

$$K_P = K_{P0} + \Delta K_P \tag{9}$$

$$K_I = K_{I0} + \Delta K_I \tag{10}$$

$$K_D = K_{D0} + \Delta K_D \tag{11}$$

In the formula, K_{P0}, K_{I0}, K_{D0} are the initial values of PID gains, which are also the initial gain parameter values of traditional PID control. The Z-N response curve method [19] is used to calculate the initial value of PID gain, and the transfer function is constructed as follows:

$$G(s) = \frac{Ke^{-LS}}{TS + 1} \tag{12}$$

Then the initial value calculation formula of PID gain is as follows:

$$K_P = \frac{1.2 \times T}{K \times L} \tag{13}$$

$$K_I = \frac{K_P}{2 \times L} \tag{14}$$

$$K_D = \frac{K_{P \times L}}{2} \tag{15}$$

Since the controlled object of the system is a temperature-controlled heater, through system identification and actual test parameter tuning, it is concluded that its model is a first-order inertial time delay system, and its transfer function is obtained by mathematical calculation [20]:

$$G(s) = \frac{6}{1.5s + 1} e^{-s} \tag{16}$$

Through the transfer function calculated by this system, the initial value of PID gain is obtained, $K_{P0} = 0.3$, $K_{I0} = 0.15$, $K_{D0} = 0.15$, which is brought into the simulation structure model of the PID control system. The model simulation results are shown in Fig. 13.
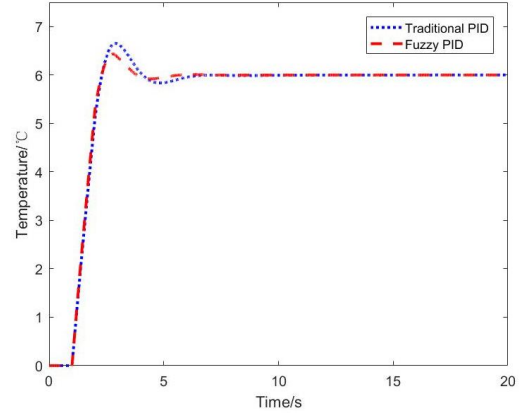
Fig. 13. PID control system simulation diagram.

The simulation result shows that, compared with the traditional PID method, the fuzzy PID control method with variable target temperature has faster response speed and lower oscillation speed. Specific performance indicators are shown in Tab. 2.

TABLE II. DIFFERENT HEATER STRUCTURE PERFORMANCE COMPARISON TABLE

Controller Category	Overshoot /%	Adjustment Time /s	Rise Time /s
Fuzzy PID	7.227	3.446	1.286
Traditional PID	10.928	4.727	1.312

It can be seen from the above table that the fuzzy PID closed-loop control method with variable parameters has the advantages of small overshoot range and high corresponding speed, which can make the system reach equilibrium in shorter time.

Humidity sensors and their heaters will encounter a certain degree of interference in a rapidly changing environment. Compared with the open-loop control method, the fuzzy PID closed-loop control method proposed in this paper also has stronger anti-interference ability. The interference simulations were performed on the two methods respectively, and the simulation results are shown in Fig. 14 and Fig. 15.
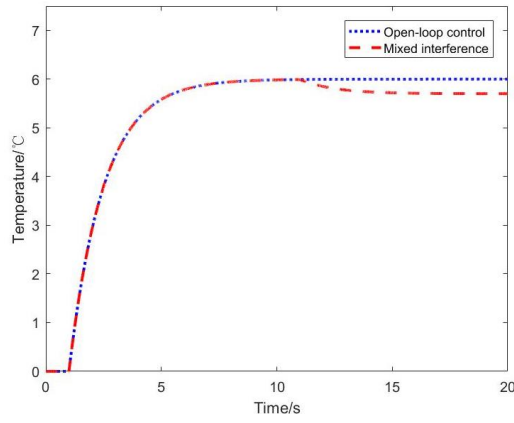
Fig. 14. Open loop control interference simulation diagram.
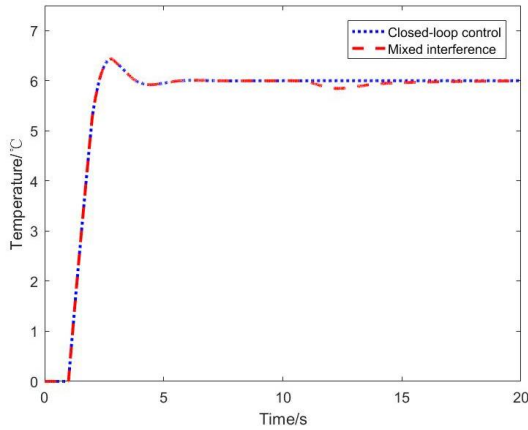
Fig. 15. PID closed-loop control interference simulation diagram.

It can be clearly seen from the figure that when the open-loop system introduces interference, the system will not reach the target temperature and cannot recover over time. The fuzzy PID closed-loop control method used in this article will adjust the interference inside the system after the interference is introduced, and restore the system to a stable state immediately, with strong environmental adaptability and anti-interference ability.

In addition to the above advantages, the fuzzy PID closed-loop control method proposed in this paper also has the performance of variable target parameters. Through the temperature data monitored by the ambient temperature sensor, the target temperature parameters are revised in real time to realize the self-tuning of the target parameters. The simulation results are shown in Fig. 16.
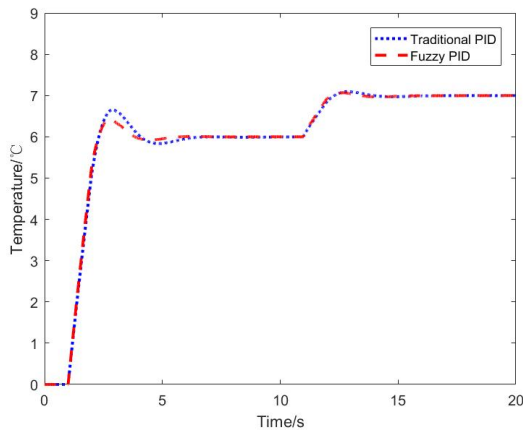
Fig. 16. Variable target temperature PID control simulation diagram.

The results show that, compared with traditional PID control, fuzzy PID control will respond to adjustments immediately after the system has experienced changes in the parameters of the target temperature, and quickly make the system return to the equilibrium state, while the traditional PID method requires longer adjustment time. It can be seen that the variable parameter fuzzy PID closed-loop control method proposed in this paper is far better than the open-loop control and the traditional PID control method. While reducing the error, it reduces the heating power consumption and meets the heating demand of the humidity sensor.

IV. CONCLUDING REMARKS

This paper proposes a controllable heating humidity sensor heating system design method for the defects of humidity sensor in high humidity and low temperature environment. The structure of heater was studied by finite element analysis. The designed structure reduces the heating power consumption to $5 \times 10^6 W/m^3$, at the same time, the temperature distribution is more even, the difference between high and low temperature is only $0.0878℃$, and the coverage is wider, reaching $4.7mm$. Simultaneously, the position distribution of the temperature sensor is guided by the temperature distribution of the heater. In order to further reduce the power consumption of the heater, the fuzzy PID closed-loop control with variable target parameters is used to replace the original open-loop heating method. The simulation result shows that, compared with the PID control method with fixed parameters, the system overshoot is reduced to 7.227%, the adjustment time is decreased to $3.446s$, and the anti-jamming ability is stronger. Future work will include designing and manufacturing the humidity sensor and its heater structure to make it more stable to serve more applications.

REFERENCES

[1] Abdurrashid Shuaibu Hassan, Vimala Juliet and C. Joshua Amrith Raj. "MEMS Based Humidity Sensor with Integration of Temperature Sensor," Materials Today: Proceedings, 2018, 5(4).
[2] Luo Yi, Yang Kun, Shi Yunbo and Shang Chunxue. "Research of radiosonde humidity sensor with temperature compensation function and experimental verification," Sensors and Actuators A: Physical, 2014, 218, pp. 49-59.
[3] Ge Verver, Masatomo Fujiwara, Pier Dolmans, Cor Becker, Paul Fortuin and Larry Miloshevich. "Performance of the Vaisala RS80A/H and RS90 Humicap Sensors and the Meteolabor 'Snow White' Chilled-Mirror Hygrometer in Paramaribo, Suriname," Journal of Atmospheric and Oceanic Technology, 2016, 23(11), pp. 1506-1518.
[4] Ferrare R A, Melfi S H, Whiteman D N, Keith Evans, Schmidlin F J and David OC. Starr. "A Comparison of Water Vapor Measurements Made by Raman Lidar and Radiosondes," Journal of Atmospheric & Oceanic Technology, 1995, 12(6), pp. 1177.

[5] Wang J, Cole H L, Carlson D J, Miller E R, Beierle K, Paukkunen A, et al. "Corrections of Humidity Measurement Errors from the Vaisala RS80 Radiosonde—Application to TOGA COARE Data," Journal of Atmospheric & Oceanic Technology, 2002, 19(7), pp. 981-1002.

[6] Voemel H, Selkirk H, Miloshevich L, Jessica V-C, Juan V, Esko K, et al. "Radiation Dry Bias of the Vaisala RS92 Humidity Sensor[J]," Journal of Atmospheric & Oceanic Technology, 2007, 24(6), pp. 953.

[7] Yang Yunxiao and Jia Li. "Experimental investigation on heat transfer coefficient during upward flow condensation of R410A in vertical smooth tubes," Journal of Thermal ence, 2015(2), pp. 155-163.

[8] Fan L S, Tai Y C and Muller R S. "IC-processed electrostatic micro-motors," IEEE International Electron Devices Meeting, 1988, pp. 666-669.

[9] FIVELAND and W. A.. "Three-dimensional radiative heat-transfer solutions by the discrete-ordinates method," Journal of Thermophysics & Heat Transfer, 1988, 2(4), pp. 309-316.

[10] Zhao Wenjie, Shi Yunbo, Luo Yi, Li Guolong and Zhou Zhen. "Design of the wind velocity sensor based on AIN thermal isolation MEMS array," Chinese Journal of entific Instrument, 2012, 33(12), pp. 2819-2824.

[11] Cun-Feng C, Sun Y R and Shui-Ming H. "Optical determination of the Boltzmann constant," Chinese Physics B, 2015, 23(005), pp. 87-92.

[12] M. Rebay, A. Arfaoui, G. Mebarki, R. Ben Maad and J. Padet. "Improvement of the pulsed photothermal technique for the measurement of the convective heat transfer coefficient," Journal of Thermal Science, 2010(04), pp. 357-363.

[13] Tang Wei, Shi Songjiao and Wang Mengxiao. "Autotuning PID control for long time-delay processes," Ifac Proceedings Volumes, 2002, 35(1), pp. 433-438.

[14] Chen G.. "Conventional and fuzzy PID controllers: An overview," Int.J.of Intelligent Control and Systems, Vol.1, 1996, pp. 235-246.

[15] WU Zhenyu, YAO Jianjun and YUE Donghai. "Fuzzy selftuning PID controllers and applications," Journal of Harbin Institute of Technology, 2009, 36(11).

[16] A. Mészáros, A. Andrášik, P. Mizsey, Z. Fonyó and V. Illeová. "Computer control of pH and DO in a laboratory fermenter using a neural network technique," Bioprocess & Biosystems Engineering, 2004, 26(5), pp. 331-340.

[17] Chen Yan, Zhang Kuangwei, Zhao Xiaoxia, Dai Yongnian, He Yuxiang and Lei Jinhui. "Design of parameter self-tuning fuzzy PID controller for vacuum metallurgy," Zhenkong Kexue yu Jishu Xuebao/Journal of Vacuum ence and Technology, 2014, 34(5), pp. 528-532.

[18] Nhon P N Q, Elamvazuthi I, Fayek H M, Parasuramanb S and Ahamed Khan M K A. "Intelligent Control of Rehabilitation Robot: Auto Tuning PID Controller with Interval Type 2 Fuzzy for DC Servomotor," Procedia Computer ence, 2014, 42, pp. 183-190.

[19] Sudheer Kasa, Sudha Ramasamy and Prabhu Ramanathan. "Hybrid fuzzy-ZN PID control based grid interfaced distribution level renewable energy source with power quality," ICCPCT, IEEE, 2015, pp. 1-7.

[20] Pokle S B, Kulat K D and Keskar A G. "MATLAB Simulation of a Fuzzy Controller for Attitude Control of a Geostationary Satellite," IETE Journal of Education, 2004, 45(4), pp. 203-209.

A Condition-Based Calibration Method for Flight Test Measuring Equipment

Renjian Feng
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
rjfeng@buaa.edu.cn

Ruoyan Xing
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
Ruoyanxing@buaa.edu.cn

Yinfeng Wu*
School of Instrumentation and Optoelectronic Engineering Chongqing
Beihang University
Beijing, China
yfwu@buaa.edu.cn

Ning Yu
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
nyu@buaa.edu.cn

Abstract—The degradation rate of the specification for the flight test measuring equipment varies with the actual working environment. The determination of the proper calibration time becomes a challenging job as the current method of calibrating test equipment in a fixed period is no longer suitable for the increasingly onerous flight test tasks. This paper proposes a method combining the degradation model and the extended Kalman particle filter algorithm to realize the state detection of the test instrument. Condition-based calibration can be achieved by this method, thereby avoiding flight test risks caused by insufficient calibration and economic waste caused by over-calibration. Finally, the equipment calibration cycle is optimized and the calibration efficiency is improved.

Keywords—flight test measuring equipment, condition-based calibration, metrological characteristic, degradation model, EKPF

I. INTRODUCTION

The flight test measuring equipment mainly completes the sensitivity, collection, recording and processing of the crucial parameters of the aircraft and various airborne equipment, and plays a role of providing parameters in the model finalization and airworthiness certification flight test. In the course of flight use, the test instrument gradually degrade due to various reasons, resulting in lower performance indicators. To ensure the accuracy of the test data, the flight test measuring equipment must be calibrated on time. The standard for calibration of flight test measuring equipment defines the specified calibration cycle as a fixed period. To make it fit the most suitable calibration interval as much as possible, the calibration cycle is generally optimized according to the performance of the instrument in the previous calibration. Gray prediction method and method based on reliability index, etc. are commonly use, but none of them have deviated from the limitation of regular calibration. The actual use, there are differences in performance of degradation process of test equipment in same model, and the performance degradation of the same instrument between several calibrations are also different. Regular inspections may cause insufficient calibration in rapid degradation of instrument and excessive calibration in slow degradation of instrument. And there is a stipulation that "Once the calibration cycle of the measuring instrument is determined, the appropriate adjustment of the calibration cycle is also allowed [1]." Therefore, the test equipment in different environments should be calibrated according to the actual situation. To reduce the waste of resources and improve calibration efficiency while ensuring the accuracy of the test equipment, calibration is only performed when necessary.

Condition-based calibration is a method to optimize the calibration cycle. In this method, the measurement process control model is established within the two calibration intervals to obtain optimally estimation of the state of the automatic test instrument. It is compared with the standard threshold to determine that the instrument is in which working stage from normal to failure, to determine the calibration time of the instrument [2][3]. By combining the degradation model and the extended Kalman particle filter(EKPF) algorithm, this paper proposes a condition-based calibration method for the flight testing measurement instrument, which is used to solve the problem of insufficient calibration and over-calibration in regular calibration cycle, to obtain the optimal solution of economic benefits and work efficiency, and improve the efficiency of flight test.

II. CONDITION-BASED CALIBRATION METHOD COMBINING THE DEGRADATION MODEL AND EKPF

In the condition-based calibration method, the main task is to optimally estimating the metrological characteristics of the test equipment. In order to improve the accuracy of the optimal estimation, it is necessary to make full use of historical calibration data and flight test information which include ground data before the flight and various parameters are collected during the flight. Therefore, a condition-based calibration method driven by test data was designed. The algorithm structure is shown in Figure 1.The steps are:

1. On the basis of historical degradation data , a test instrument state degradation model is established combined with the principle of test instrument degradation;

2. Monitor the ground data of the flight test measuring equipment with principal component analysis(PCA) method, and initially solve the metrological characteristics of the test instrument;

3. Perform data analysis on flight data combining with ground data, and use the correlation of measuring instruments at different locations to solve the metrological characteristics of equipment;

4. Combining the degradation model and the metrological

characteristics of equipment, the optimal estimation of the metrological characteristics of the test equipment is performed with the EKPF algorithm. It is compared with the standard threshold to realize the condition-based calibration and optimize the calibration cycle.
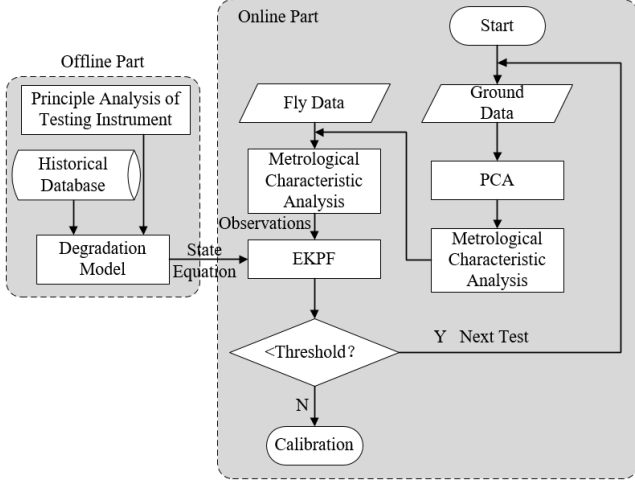
Fig 1. Signal of the condition-based calibration method

A. Degradation Model of Test Equipment

To analyze the degradation process of the instrument state, a degradation model needs to be established at first [4]. The degradation model based on the Wiener process is widely used because of its good mathematical properties and clear physical explanation.

The metrological characteristics of the test equipment change slowly with time, and its degradation process is a random process. The degradation model based on the Wiener process of test equipment is:

$$X(t) = X(0) + \int_0^T r[w_1(t),...,w_n(t)] + \sigma B(t) \quad (1)$$

where X(t) is the degradation amount of the equipment at time t, X(0) is the degradation quantity of the test equipment at the initial time, $r[w_1(t),\cdots,w_n(t)]$ is the performance degradation rate function of the test equipment, σ is the diffusion parameter, B(t) is the standard Brownian motion.

The metrological characteristics of test equipment are affected by both its own degradation and external factors. In the performance degradation rate function $r[w_1(t),\cdots,w_n(t)]$ of test equipment, $w_1(t),\cdots,w_n(t)$ respectively represent environmental factors such as temperature, humidity, and salt spray concentration at time t. According to the multi-stress generalized Arrhenius model form [5], the degradation rate function of test equipment is established:

$$r[w_1(t),...,w_n(t)] = b_0 + b_1 \square w_1(t)^p + ... + b_n \square w_n(t)^p \quad p = \pm 1 \quad (2)$$

Both the independent and interactive terms of environmental stress will affect the degradation rate of test equipment. However, through most experimental observation results analysis, independent terms are the main terms, and the interactive terms are secondary generally. And in the case of many environmental factors, the calculation of interactive terms is extremely complicated. In order to reduce the number of parameters to be estimated, only independent terms are retained in formula (2), b_0、$\cdots$、b_n are the parameters to be estimated, and b_0 is the embodiment of the equipment's own degradation factor.

Because of the independent incremental characteristic of the Brownian motion process, accumulate and approximate the integral part of the degradation model, and then estimate the estimated parameters b_0、$\cdots$、b_n and the diffusion parameter σ. By simplifying the model, the degradation model is approximately:

$$X(t) \approx X(0) + \sum_0^m r[w_1(t),...,w_n(t)]\Delta t_i + \sigma B(t) \quad (3)$$

In the formula, m is the number of accumulated observations of degraded variables from time 0 to time T, Δt_i is the observation time interval,$\Delta t_i = t_i - t_{i-1}$,$w_k(t_i)$ is the environmental stress in the time interval $[t_{i-1}, t_i]$, $k = 1,2,\cdots,n$.

According to the independent incremental characteristic of Brownian motion, there are:

$$\Delta X(t_i) = X(t_i) - (t_{i-1}) \quad (4)$$

$$\Delta X(t_i) \square N(r[w_1(t),...,w_n(t)], \sigma^2 \Delta t_i) \quad (5)$$

$$\frac{\Delta X(t_i)}{\Delta t_i} \approx r[w_1(t),...,w_n(t)] \quad (6)$$

The $\Delta X(t_i)/\Delta t_i$ result is obtained by using the degradation data of the flight test measuring equipment, and the estimated parameters in the degradation rate function are estimated through regression analysis.

The diffusion parameter σ in the degradation model is estimated by the maximum likelihood estimation method. First, the cumulative effect in the data is eliminated, and there are:

$$\Delta H(t_i) = X(t_i) - X(0) - \sum_{i=1}^m r[w_1(t),...,w_n(t)]\Delta t_i = \sigma B(t) \quad (7)$$

The likelihood function of the diffusion parameter σ can be obtained as:

$$L(\sigma) = \prod_{i=1}^m \frac{1}{\sigma\sqrt{2\pi\Delta t_i}} \exp[-\frac{\Delta H(t_i)^2}{2\sigma^2 \Delta t_i}] \quad (8)$$

The estimated value of the diffusion parameter σ is:

$$\hat{\sigma} = \sqrt{\frac{1}{m}\sum_{i=1}^m [\frac{\Delta H(t_i)^2}{\Delta t_i}]} \quad (9)$$

B. Analysis of Metrological Characteristics

In the laboratory environment, a higher-level instrument is required to provide the reference value, and such conditions are not available on the flight test vehicle when analyzing the metrological characteristics of the equipment to be calibrated.

Therefore, it is necessary to use some of the more stable device parameters on the flight test vehicle to convert them into reference values that can be used to analyze the metrological characteristics of the test equipment.

In the complete flight test, the test equipment will obtain the ground data before takeoff and the flight data after takeoff. Because of the measured data dynamic changes during flight, it is difficult to directly provide a good reference value. Therefore, it is necessary to use ground data to provide a steady-state reference value in order to obtain preliminary the metrological characteristics, and analyze the flight data to obtain further metrological characteristics in next step.

When using the ground data for reference, changes in the state of the device under test and degradation of the

performance of the test instrument will cause significant differences between the test data and the data in normal state. Therefore, it is necessary to check whether the state of the test object is normal firstly, and then eliminate the data fluctuation caused by the abnormal state of the test object. Consequently, the state of the testing machine needs to be monitored. The Model-Based method, method based on deep learning, and data-driven approach, etc. are condition monitoring methods commonly used. There are a large number of highly correlated process variables in the flight test experiment, which are essential to ensure the normal conduct of the flight test and the reliability of the test machine. It is difficult to monitor so many process variables simultaneously with the general method. In addition, when some instruments fail, the mean and variance of the process variables are basically unchanged but the correlation between the process variables changes. These cannot be accurately identified in the traditional monitoring system while the PCA method provides support for solving such problems.

Use the PCA method for data processing to monitor the status of the tested parts of the flight test vehicle: First, establish the principal component offline model with the PCA algorithm, and the sample data came from the laboratory data and historical data of the flight test vehicle[6]. Then select the square prediction error (SPE) statistic as the failure detection index, and calculate the control limit of the SPE statistic. Next bring the ground steady-state data obtained by the test equipment into the offline principal component model and calculate the SPE statistics. When the SPE statistics exceed the control limit, an early warning will be issued to indicate that the tested part may be faulty, otherwise the tested part is considered to be in normal state. Take the theoretical value of the measured part parameter as the reference value, analyze and obtain the metrological characteristics of the test instrument preliminarily.

After the flight test started, most of the measured values of the test equipment fluctuated greatly, and other measuring instruments with relevant measured values were used to provide references for the calibrated instruments. The measurement data of other relevant measuring instruments are corrected by the above-mentioned initial metrological characteristics obtained by using ground data, and then converted into the corresponding measurement values of the test instrument to be calibrated. According to the correlation size and stability level of different instruments, the measured values are weighted and synthesized into the reference value of the equipment under calibrated.

C. Optimal State Estimation Obtained by EKPF

When the test equipment is calibrated and then installed on the flight test vehicle, this moment is recorded as the initial time. At this time, the metrological characteristics of the test equipment are known. Each subsequent test and long-term storage will cause the performance of the test equipment to degrade. Combining the above-mentioned degradation model and the analysis of the metrological characteristics, iteratively predict the metrological characteristics of the test instrument at this moment, and the filtering method can be used.

The EKPF algorithm is one of the most commonly used method to solve the problem of nonlinear state variable estimation [7]. Particle filtering gets rid of the constraint that random variables must satisfy Gaussian distribution when solving nonlinear filtering problems, and can be applied to any nonlinear non-Gaussian random system. The recommended

value of the density distribution in the particle filter will affect the performance of the particle filter. In order to avoid the particle depletion problem in particle filtering, EFK is used as the recommended distribution to realize the estimation of the metrological characteristics obtained by the EKPF algorithm.

Use the aforementioned degradation model as the discrete state model of the measuring instrument. The metrological characteristics of the instrument are used as a system variable, and the metrological characteristics obtained by solving the instrument data are used as an observation variable, then there is the state equation:

$$x_{k+1} = x_k + r[w_1(t),...,w_n(t)] + \frac{\sigma}{m} B(t) + v_k \qquad (10)$$

Observation equation:

$$z_k = x_k + n_k \qquad (11)$$

Among them: x_k is the metrological characteristics at time k, z_k is the observed value at time k, v_k and n_k are system noise and measurement noise, respectively.
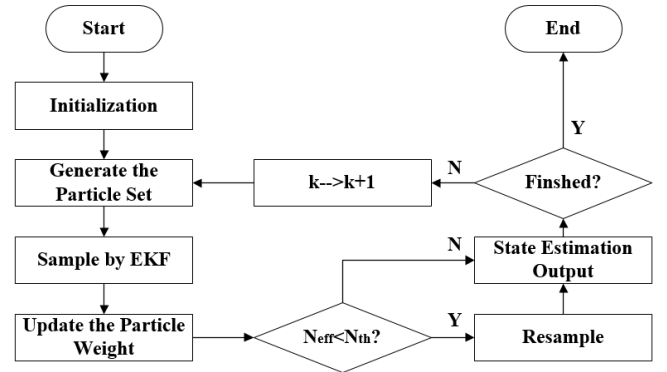
Fig 2. EKPF algorithm flow chart

According to the state equation and observation equation of the metrological characteristics of the test instrument, the algorithm flow of using EKPF to realize the state estimation is shown in Figure 2. The specific algorithm steps are as follows:

1. Initialization

Generate the particle $x_0^i, i = 1, ..., N, k = 0$ from the known prior probability density $p(x_0)$;

2. Importance sampling:

Use the EKF algorithm to sample $x_k^i \sim p(x_k|x_{k-1}^i)$, update the particle weight $w_k^i = w_{k-1}^i p(z_k|x_k^i)$, and normalize the weights $\widetilde{w}_k^i = w_k^i / \sum_{i=1}^N w_k^i$;

3. Calculate $N_{eff} = 1/ \sum_{i=1}^N (w_k^i)^2$.

If N_{eff} is less than the preset threshold N_{th}, resample particle:

Resample x_k^i to generate a new particle set $\widetilde{x}_k^i$, and set the weight of all particles to $\widetilde{w}_k^i = 1/N$;

4. State estimation output:

If there is no resampling: $\hat{x}_k \approx \sum_{i=1}^N x_k^i \widetilde{w}_k^i$;

If there is resampling: $\hat{x}_k \approx (\sum_{i=1}^N x_k^i)/N$;

5. Set k=k+1 and return to step 2.

III. SIMULATION TEST AND EFFECT ANALYSIS

In order to verify the feasibility of the method proposed above, a simulation test was performed on an airborne pipeline pressure test instrument [8].

According to the laboratory data and historical calibration data of the stress test equipment, the degradation model was established as:

$$X(t) = X(0) + \sum_{i=1}^{m} r[w_1(t),...,w_n(t)]\Delta t_i + \sigma B(t) \quad (12)$$

$$r = b_0 + b_1 \square w_1(t) + b_2 \square w_2(t)^{-1} + b_3 \square w_3(t) \quad (13)$$

In the formula: b_0, b_1, b_2, b_3 are the constant degradation parameters, temperature, humidity and salt spray degradation parameters of the metrological characteristics, respectively, $w_1(t), w_2(t), w_3(t)$ are temperature, humidity and salt spray concentration respectively.

Before the flight test vehicle started, the pressure in the pipeline is relatively stable. After the testing aircraft started, with different instructions from the fuel control instrument, the flow rate and flow speed of the fuel pumped by the fuel pump fluctuate greatly, and the fuel pressure in the fuel pipeline changes more. It is complicated, and even transient processes such as pressure shocks may occur. The pressure change of the pipeline is shown in Figure 3.
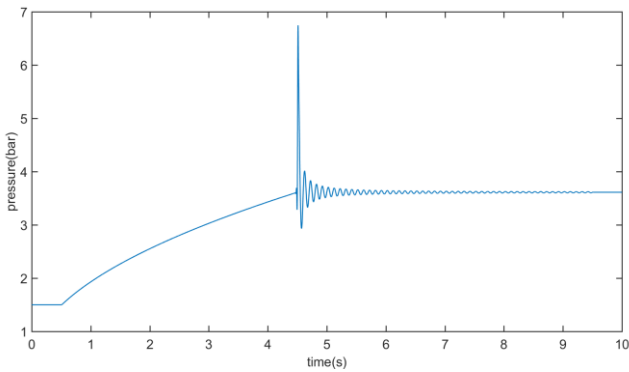
Fig 3.   Simulation of an aircraft fuel line pressure

Process the data before and after startup orderly. The ground data is preprocessed by PCA algorithm, the SPE statistics are solved, and whether the pipeline is faulty. When there is no leakage in the pipeline, the SPE statistics are within the control limit, and the ground data will be used to obtain the metrological characteristics initially. When the flight data are collected, the metrological characteristics of the measuring instrument were monitored by using the correlation between the measured pressure values.

Taking the pressure relationship along the path in a common straight pipe as an example for simulation, there is a phenomenon called pressure drop in a straight pipe, which show that the pressure value measured at the back position along the liquid flow direction is lower than the front position. The value of pressure drop has a certain functional relationship with the pipe length, diameter, material, fluid velocity, type and some other parameters. These pressure values measured at different positions are highly correlated, which are used in piping system design. As shown in Figure 4, the pressure testing instruments is installed at multiple positions on the fuel pipeline. Positions 1 and 2 are

symmetrical with respect to the pipeline. Positions 3 and 4 are behind positions 1 and 2. The corresponding instruments are marked as No. 1~4. Ignoring the influence of the uneven distribution of the liquid along the pipe diameter, the pressure generated by the liquid in the pipeline measured at position 1 and 2 are the same, and the pressure values measured at position 3 and 4 have the certain drop and delay compared with that at position 1. After the values measured at positions 2, 3, and 4 were compensated by the ground data, the data at positions 3 and 4 were converted into the equivalent values at position 1 through the inverse function.
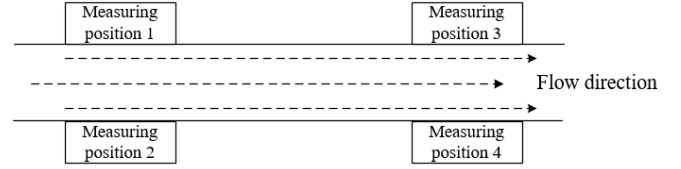
Fig 4.   Pipeline pressure measurement position indication

The converted data and the data at position 2 were weighted and combined into a reference pressure value at position 1. Then obtained the metrological characteristics of the measuring instrument at position 1 by analyzing data.

Estimated values were obtained by the EKPF algorithm, for which equations (12) and (13) served as the state transition equations, and the observation values came from the metrological characteristic settlement values obtained with the method as mentioned above. The estimation result is shown in Figure 5. It can be seen from the simulation results of Fig. 5 that the estimated value of the metrological characteristics of measurement equipment effectively conforms to the theoretical value and that the condition-based calibration method combining the degradation model and EKPF achieves the tracking estimation of the change in the measurement equipment's metrological characteristics.
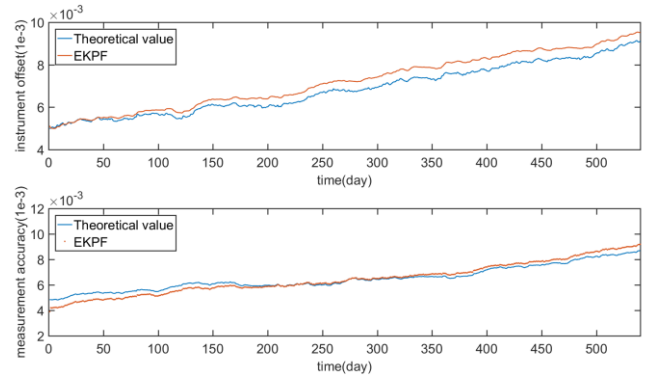
Fig 5.   Estimation results of test equipment's metrological characteristics

Take quarter as the time node to evaluate the estimation results. The results are shown in Table 1 and Table 2. The estimated and theoretical values are all relative value.

For this equipment, assume that the calibration threshold of measurement accuracy is 0.1% and the calibration threshold of instrument bias is 0.1%. Set one year as the calibration period. It can be seen from the simulation results of Table 1 and Table 2 that the degradation rate of the instrument was much slower than expected, and the measurement characteristics did not exceed the threshold until the sixth quarter. Compared with the regular calibration method, this method requires the instrument to be calibrated in the sixth quarter, avoiding over-calibration. If the

requirements for the measurement characteristics of the test equipment are increased and the instrument bias of 0.7% is

Table 1. COMPARISON OF ESTIMATED AND THEORETICAL VALUES OF MEASUREMENT ACCURACY

Time(day)	Theoretical accuracy (10^{-3})	Estimated accuracy (10^{-3})	error (10^{-3})
0	4.95	5.16	0.21
90	5.41	5.53	0.12
180	5.93	6.11	0.18
270	6.62	6.78	0.16
360	7.03	7.24	0.21
450	7.91	8.05	0.14
540	9.44	9.53	0.09

Table 1. COMPARISON OF ESTIMATED AND THEORETICAL VALUES OF INSTRUMENT BIAS

Time(day)	Theoretical accuracy (10^{-3})	Estimated accuracy (10^{-3})	error (10^{-3})
0	4.81	4.20	-0.39
90	5.60	5.23	-0.37
180	6.56	6.30	-0.26
270	7.09	7.25	0.16
360	7.96	8.33	0.37
450	9.51	9.97	0.46
540	10.02	10.67	0.65

used as the calibration threshold, the equipment should be calibrated in the third quarter. This method not only can avoid insufficient calibration but also more flexible to use when compared with the one-year calibration cycle. Though both the degradation conditions and the requirements for the measurement value characteristics of the instrument are different in varied environments, it`s feasible to estimate the metrological characteristics of the instrument. Thereby, the calibration period can be adjusted more accurately, and the efficiency of the calibration work can be improved.

IV. CONCLUSIONS

Aiming at the problem of insufficient calibration or over-calibration caused by the annual calibration cycle of the flight test measuring equipment in China, a condition-based calibration method combining the degradation process and EKPF was designed. A simulation experiment was carried out on the airborne fuel line pressure test equipment as an example to verify the method. The experimental results showed that the method can accurately estimate the metrological characteristics of the flight test measurement equipment to achieve condition-based calibration. The calibration efficiency can be improved and the cost of the calibration work can be reduced.

Condition-based calibration is the development trend of the method to determine the optimal calibration cycle of the instrument. In this method proposed in this paper, laboratory data and historical data of test equipment degradation are required. It is not easy to obtain data like these in practical applications. With only a small amount of historical degradation data and laboratory data, it is a difficult point for future research to make more accurate predictions for the degradation of the metrological characteristics of flight test measuring equipment.

REFERENCES

[1] "JJF 1139-2005 Principle and Method for Determination Verification Period of Measuring Instruments". Beijing, State General Administration of the People's Republic of China for Quality Supervision and Inspection and Quarantine 2005.12

[2] Zuonian. Guan, Naixiang. Wang, and Ning Xu, "Analysis of angle accuracy of airborne photoelectric platform based on Monte Carlo simulation", Journal of Electronic Measurement and Instrumentation, vol.29, pp.447-453,Mar,2015.

[3] Zhongjie. Wang, "Calibration cycle of flight parameters test system optimization management strategy research", Foreign Electronic Measurement Technology,vol. 36, pp.91-93,Mar,2017.

[4] Shengjin. Tang, Chuanqiang. Yu, Xue. Wang, Xiaosong. Guo, and Xiaosheng.Si, "Remaining Useful Life Prediction of Lithium-Ion Batteries Based on the Wiener Process with Measurement Error". Energies, vol.7, pp.520-547,Jan,2014.

[5] Zhen. Lin, Tongmin. Jiang, Yongsheng. Cheng, and Bin. Hu. "Study on Arrhenius relationship". Electronic Product Reliability and Environmental Testing, vol.23,pp.12-14,Jun,2005.

[6] Lin H. Shang, "A survey of functional principal component analysis", AStA Advances in Statistical Analysis, vol.98,pp.121-142,Feb,2014.

[7] A. Doucet, S. Dodsill and C. Andrieu. "On sequential Monte Carlo sampling methods for Bayesian filtering", Stats & Computing, vol.10,pp.197-208,Sep.2000.

[8] Zehai. Gao, Cunbao. Ma, and Dong. Song, "Research of Aircraft Fuel Feeding System Based on Flowmaster Simulation". Computer Measurement & Control, vol.23,pp.1703-1709,May.2015.

Design of the Self-test Structure of QVBA

Shengshou Lin
School of Instrument Science and Engineering
Southeast University
Nanjing, China
15858023900@163.com

Jinxing Liang*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
j-liang@seu.edu.cn

Abstract—In this paper, the overall structure of QVBA and its stress and deformation mechanism are specifically analyzed in details. The pull-in phenomenon is considered, and the general analytical methodology of pull-in is investigated. An effective theoretical modal of pull-in effect for QVBA is proposed and demonstrated to be correct and precise by taking advantage of the finite element method. Meanwhile, the deformation of isolating element is proved to be non-negligible. Mainly considering the pull-in effect, a self-test structure is designed with the initial spacing of 60μm, and the simulating acceleration range is optimized up to ±4g.

Keywords—QVBA, pull-in, self-test

I. INTRODUCTION

Reliability is becoming more and more important for MEMS accelerometer, especially for military applications. There are some inducing factors which will reduce the sensor's accuracy and even cause its failure. For example, machining errors or surface defects are likely to occur during micro machining, such as poor roughness and residual mechanical stress. These factors will reduce the performance and the service life of the instrument and cause unnecessary losses [1]. In fact, the methods of testing the performance of accelerometer after fabrication have been thoroughly studied in past decades, which ensure its accuracy and reduce the loss [2-5]. Unfortunately, these testing methods are not applicable in all cases. In some integrated control systems, inertial devices and other devices are embedded together in mechanical systems. In order to ensure that it can continue working normally, the inertial devices should be tested after a period of use regularly. However, the disassembly is obviously not an appropriate method. The operation is not simple, and it may also damage the installation accuracy of the control systems and leads to recalibration. Therefore, a built-in self-testable function is desired, which can provides the on-line test.

The self-test of the micro-accelerometer, that is, by applying a suitable excitation on the sensitive element, the inertial force(due to acceleration) can be imitated. Then, the performance of the accelerometer can be checked according to the output. The common excitation methods include electromagnetic way, thermal way and electrostatic way etc. The structure of electromagnetic excitation is generally suitable for vibration-excitation accelerometer, and the applied range is limited relatively[6]. The thermal excitation uses thermal deformation to generate the driving force, and the magnitude of the driving force is sufficient. However, the thermally excited elements are very sensitive to temperature, and small temperature change is also easy to bring about deviation from the result. In addition, thermal excitation is a gradual process, so it has low efficiency both in excitation and recovery [7]; As for electrostatic excitation, the force is distributed in planar mode, quite similar to the gravity that the distribution is body mode. Besides, its excitation efficiency is high and the excitation structure is easy to design.

Consequently, the electrostatic excitation is more suitable for the self-test of accelerometer.

Quartz vibrating beam accelerometer (QVBA) is a solid-state sensor based on the resonance principle of in-plane bending vibration mode (Fig. 1). Through the proof-mass, the input acceleration information can be converted to an inertial force and acts on the resonator. Due to the force-frequency effect, the vibration frequency of resonator will vary, and the acceleration value can be measured by detecting the frequency difference of the resonator.

In this paper, we propose a self-test structure, using electrostatic excitation, to realize the self-testable function of QVBA (Fig. 2). Specifically, along the detection direction, a parallel flat plate is placed on each side of the accelerometer mass. A conductive film is plated on the surface of the mass and the corresponding mass area on the inside of the two plates, and a voltage is applied to generate a certain electrostatic force.

However, the electrostatic force results in nonlinear dynamics, due to the known "pull-in effect". Namely, the mass is caused to collapse on the plate due to the electrostatic force increasing more rapidly than the restoring force. Moreover, the pull-in effect limits the travel ranges of mass and leads to the limit of the self-test range of accelerometer.

Since Nathanson et al. [8] and Taylor's [9] pioneering attempt on the pull-in effect in 1960s, many researchers have studied the pull-in phenomenon. Although a lot of effort is being spent on the pull-in effect of some accelerometer structures, including the simplest spring-mass structure, cantilever beam structure [10], symmetric beam island structure [11], micro Mirror structure [12], etc. The consideration of pull-in phenomenon in self-test structure of QVBA has yet to be developed well. Currently, some researches have conducted some simulation and qualitative analysis of the self-test design of QVBA. However, to ensure the safety and accuracy of the self-test design, it is necessary to obtain a more comprehensive understanding of the pull-in in the QVBA.

Herein, we analyze the structural characteristics of QVBA and characterize its deformation and stress mechanism. Importantly, we investigate the pull-in effect in the QVBA and an overall analytical model of pull-in effect is proposed theoretically. Furthermore, taking advantage of finite element method, we simulate the self-test state and the pull-in phenomenon for verifying the proposed model. Eventually the details of the self-test structure are discussed, and the design is completed.
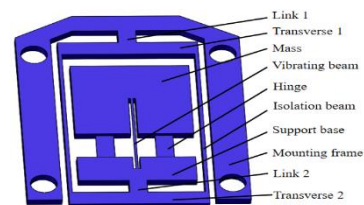
Fig. 1. 3D structure diagram of QVBA

---

* Jinxing Liang is the corresponding author. (e-mail: j-liang@seu.edu.cn)

(a) Overview of the self-test structure   (b) cross-sectional view
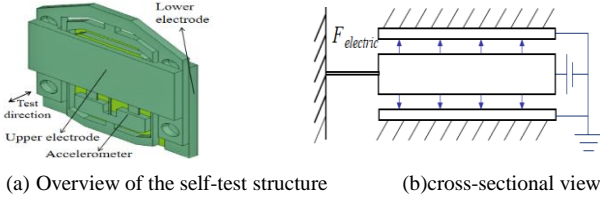
Fig. 2. The diagram of self-test structure of QVBA

II. THEORETICAL ANALYSIS

A. General Analytical Methodology of Pull-in

In general sensors, the electrode plates of electrostatic excitation is usually in an parallel manner, and the pull-in phenomenon has been investigated thoroughly. However, in the case of QVBA, the structure of electrostatic excitation is non-parallel plates alike. Thus it is necessary to focus on an easiest modal to find out a general analytical methodology of pull-in.

The capacitor-mass-spring model, first proposed by Nathanson et al. and Taylor, is the easiest modal to be studied. Fig. 3 shows the details about the modal.

Easily to be known that the mechanical force and electrical force can be presented by equations (1) and (2) respectively:

$$F_m = -k_m \cdot x \tag{1}$$

$$F_e = \frac{\varepsilon S U^2}{2(x_0 - x)^2} \tag{2}$$

wherein, k_m is the stiffness of spring, x_0 is the initial spacing of electrode plates. And Fig. 4 shows the image of $|F_m|$ and $|F_e|$ with respect to the degree of freedom.

It can be seen from the statics analysis that when the two curves have an intersection point, it indicates that the system is under force-balanced. Otherwise the electrostatic force is greater than the elastic force, and the mass continues to move downward until it is attracted to the fixed electrode.

At the intersection point, there are two vital features. First, the electrostatic force is equal to the elastic force, namely, the forces are balanced; second, the slope of the electrostatic force curve and the elastic force line are equal at this point. Here the slope of the electrostatic force curve is defined as electric stiffness. Therefore, the characteristics of pull-in phenomenon at critical state can be summarized as the following equations (3):

$$\begin{cases} \sum \vec{F}_i = 0 \\ \sum k_i = 0 \end{cases} \tag{3}$$

Fig. 3. Capacitor-mass-spring model
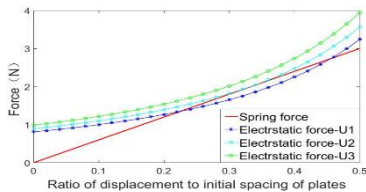
Fig. 4. Relationship between $|F_m|$ and $|F_e|$

that is, the forces of system are balanced, and the mechanical stiffness of overall structure is equal to the equivalent electric stiffness. That's exactly the analytical methodology of pull-in effect.

B. Theoretical Analysis of Sensing Element

As shown in the Fig. 1, besides the mounting frame, the structure of accelerometer can be divided into two parts: the sensing element and the isolating element. The former includes mass, vibrating beam and hinge; while the latter includes support base, link 1, link 2, transverse 1 and transverse 2. Since the stress mechanism of sensing element is very complex, it's impossible to study the subelements separately. Thus section 2.2 briefly surveys the stress mechanism of the sensing element, and the deformation mechanism of its subelements is investigated.

Yang Ting once conducted modeling and stress analysis on the sensing element of the accelerometer [13], but in his analysis there is no detailed process and the model is too idealized. The sensing element will be analyzed in detail below.

Considering that the mass is subjected to a distributed force, the vibration beam and the end of the hinge are subjected to a quasi-directed force. Moreover, it is assumed that the tilted angle of the mass is the same as the tilted angle of the hinge end, accounted that the hinges are wide enough and in the middle of the thickness of the mass, and are arranged symmetrically. In addition, some forces can be reasonably simplified according to material mechanics. The simplified force analysis diagram is illustrated in Fig. 5.

Meanwhile, the force balance equation is derived as the following equations (4):

$$\begin{cases} \sum F_{y'} = 0, & 2F_{l1} - F_b = 0 \\ \sum F_{x'} = 0, & 2F_{l2} - F_e = 0 \\ \sum M = 0, & F_b \cdot l_2 - F_e \cdot l_1 - 2M = 0 \end{cases} \tag{4}$$

wherein, F_b is the axial force of vibrating beam; F_{l1}, F_{l2} are the axial and radial force of flexible hinge respectively; M is the moment of the end of beam; F_e is the electrostatic force; l_1 is the distance from the equivalent electrostatic force application point to the end of the hinge in the y' direction; l_2 is the distance from the center of the end of the vibration beam to the center of the hinge end in the x' direction.

To have a more comprehensive understanding of this force balance equation, it is necessary to analyse the unknown terms. The following will deduce the representation of the forces respectively and try to analyze the relationship between them quantitatively.
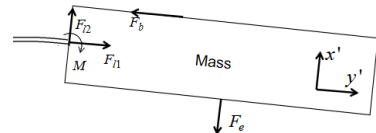
Fig. 5. Diagram of simplified force analysis

Firstly, based on the principle of bending deformation superposition of material mechanics, the deflection and obliquity of the hinge endpoint are the result of the co-action of the shear F_{l2} and the moment M, and can be given respectively by the equations (5) and (6):

$$w = w_F + w_M = (\delta_1 + \delta_2 l_1) \cdot \frac{F_e}{2} - \delta_2 l_2 \cdot \frac{F_b}{2} \qquad (5)$$

$$\theta = \theta_F + \theta_M = \left(\frac{3}{2l}\delta_1 + \frac{2l_1}{l}\delta_2\right) \cdot \frac{F_e}{2} - \frac{2l_2}{l}\delta_2 \cdot \frac{F_b}{2} \qquad (6)$$

where w_F, θ_F are the deflection and angle under shear F_{l2}, respectively; while w_M、θ_M are the deflection and angle under moment M; l is the length of the hinge, and their relations is given by following equations (7):

$$\begin{cases} w_F = \delta_1 F, & \delta_1 = \frac{l^3}{3EI} \\ w_M = -\delta_2 M, & \delta_2 = \frac{l^2}{2EI} \\ \theta_F = \frac{3}{2l} w_F = \frac{3}{2l}\delta_1 F \\ \theta_M = \frac{2}{l} w_M = -\frac{2}{l}\delta_2 M \end{cases} \qquad (7)$$

Second, considering that the vibrating beam is mainly subjected to axial force F_b, its elongation needs to be described. The detail is shown in Fig. 6.

The deformation process of the vibrating beam can be regarded as that the triangle formed by the vibrating beam on the structure is first rotated around the point O by an angle θ, and then translated in x-axis direction by a distance, which is equal to the deflection of the hinge endpoint. Compared with the length of vibrating beam, its deformation in x-axis direction is negligible. Consequently, the elongation of vibrating beam is approximately equal to the tiny radian of triangle turning around O point, and may be derived as the equation (8):

$$\Delta l = l_2 \cdot \theta \qquad (8)$$

thus, the axial force of vibrating beam is given by the equation (9):

$$F_b = \frac{\Delta l}{l_b} EA_b = \frac{l_2}{l_b} EA_b \cdot \theta \qquad (9)$$

wherein, E, A_b, l_b are the elastic modulus, cross-sectional area and length of the vibrating beam.

Thirdly, for the accuracy of the model, an analytical model of the electrostatic force between non-parallel plates will be proposed and a non-parallel plate model is established as shown in Fig. 7.
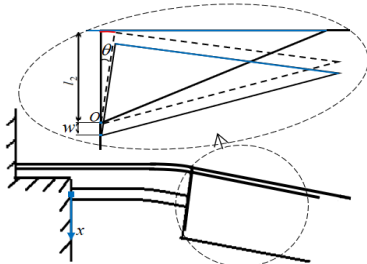
Fig. 6. Details of Vibration beam deformation
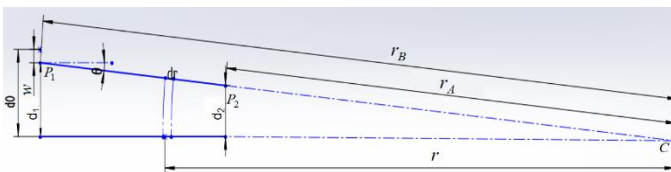
Fig. 7. Non-parallel plate model

The extension of the two plates intersect at point C and the angle is θ. Taking C as the center of the circle, when taking an arc-shaped element dr, the expression of the electric field energy element can be obtained as the equation (10):

$$dW = \frac{1}{2} dC \cdot U^2 = \frac{1}{2} \frac{\varepsilon b dr}{r\theta} \cdot U^2 \qquad (10)$$

where U is the applied voltage; ε is the dielectric constant; b is the width of mass. In accordance with the geometric relationship, the electrostatic force is derived by performing integral calculation by the equation (11):

$$F_e = \int dF = -\int_{r_A}^{r_B} \frac{\partial(dW)}{\partial(r\theta)} = -\frac{1}{2}\frac{\varepsilon b U^2}{\theta^2}\left(\frac{1}{r_A} - \frac{1}{r_B}\right)$$
$$= -\frac{\varepsilon ab U^2}{2} \cdot \frac{1}{\frac{d_2}{2\sin\frac{\theta}{2}}\left(\frac{d_2}{2\sin\frac{\theta}{2}} + a\right) \cdot \theta^2} \qquad (11)$$

Taking advantage of the equivalent infinitesimal of θ, the electrostatic force may be rewritten by the equation (12):

$$F_e = -\frac{\varepsilon ab U^2}{2d_2\left(2a\cdot\sin\frac{\theta}{2} + d_2\right)} = \frac{\varepsilon ab U^2}{2(d_0 - x)(d_0 - x - a\theta)} \qquad (12)$$

Consequently, with the above results, the deflection of the hinge endpoint relative to support base is determined by the following equation (13):

$$x = \delta \cdot F_e \qquad (13)$$

where $\delta = \frac{1}{2}\left[(\delta_1 + \delta_2 l_1) - \frac{\delta_2 l_2}{2} \cdot \frac{\frac{l_2}{l_b}EA_b \cdot \left(\frac{3}{2l}\delta_1 + \frac{2l_1}{l}\delta_2\right)}{1 + \frac{l_2^2}{ll_b}\delta_2 EA_b}\right]$

$\delta = \frac{1}{2}\left[(\delta_1 + \delta_2 l_1) - \frac{\delta_2 l_2}{2} \cdot \frac{\frac{l_2}{l_b}EA_b\left(\frac{3}{2l}\delta_1 + \frac{2l_1}{1}\delta_2\right)}{1 + \frac{l_2^2}{ll_b}\delta_2 EA_b}\right]$, and its reciprocal is the stiffness of overall structure (in the hinge endpoint). Meanwhile, the obliquity of the hinge endpoint is derived as the equation (14):

$$\theta = \eta \cdot F_e \qquad (14)$$

where $\eta = \frac{1}{2} \frac{\frac{3}{2l}\delta_1 + \frac{2l_1}{l}\delta_2}{1 + \frac{l_2^2}{ll_b}\delta_2 EA_b}$ $\eta = \frac{1}{2} \frac{\frac{3}{2l}\delta_1 + \frac{2l_1}{1}\delta_2}{1 + \frac{l_2^2}{ll_b}\delta_2 EA_b}$

Hence the relationship between the deflection and the obliquity is given by the equation (15):

$$\theta = \xi \cdot x \qquad (15)$$

Furthermore, the deflection of the mass endpoint relative to support base is derived as the equation (16):

$$X = x + a \cdot \sin\theta = x + a \cdot \theta = x + a \cdot \xi x = (1 + a\xi)x \qquad (16)$$

C. Theoretical Analysis of Isolating Element

Section 2.3 focuses on the isolating element, and the deformation mechanism of its subcomponents is described.

There is a slight deflection angle and displacement due to bending deformation at the link 1. Compared with the overall deformation, the deformation of the link 1 is very small, and it can be negligible. So does the link 2.

The transverse 1 is subjected to a downward force and a moment. The displacement caused by the downward force is relatively small compared with the overall situation, and it's negligible; while the torsion angle caused by a moment has a certain effect on the overall deformation, and should not be ignored. So does the transverse 2.

The torsion angle of the transverse 1 is given by the equation (17):

$$\theta_0 = \delta_{\theta 0} \cdot F_e \qquad (17)$$

where $\delta_{\theta 0} = -\mu_0 \cdot \dfrac{(l_6 - l_5) \cdot l_h}{G_{YZ} \beta h b^3}$.

wherein, μ_0 is an adjustment coefficient; G_{YZ} is shear modulus; β is a coefficient related to h/b; while h, b are the long and short sides of the cross-section of the transverse, respectively; l_5, l_h are half the width and half the length of transverse, respectively; and l_6 is the distance from the applied point of the equivalent electrostatic force to the transverse 1.

Similarly, the torsion angle of the transverse 2 is given by the equation (18):

$$\varphi = \mu \delta_\varphi \cdot F_e \qquad (18)$$

where $\delta_\varphi = \dfrac{(l_1 + l + l_3 + l_4 + l_5) l_h}{2 G_{YZ} \beta h b^3}$,

wherein, l_1 is the distance from the applied point of the equivalent electrostatic force to the hinge endpoint; l_3, l_4 are the length of the support base and the junction, respectively; $\mu(\mu = 1.25)$ is an adjustment factor added to the torsion angle, since the stress condition of the link 2 is very similar to that of the transverse 2, and their force and deformation are linearly related.

In addition, the force state of each isolation beam is given by the following equations (19):

$$\begin{cases} F_v = \dfrac{1}{2} F_e \\ M_v = \dfrac{1}{2} F_e \cdot (l_1 + l + l_3 + l_4 + l_5) \end{cases} \qquad (19)$$

Referring to the solution method of the flexible hinge above, the deflection and obliquity of each isolation beam are respectively given by the equations (20) and (21):

$$x_v = \delta_{vx} \cdot F_e \qquad (20)$$
$$\theta_v = \delta_{v\theta} \cdot F_e \qquad (21)$$

where $\delta_{vx} = \dfrac{\left(\frac{l_v}{3} - \frac{l_1 + l + l_3 + l_4 + l_5}{2} \right) l_v^2}{2 E_Y I_v}$,

$\delta_{v\theta} = \dfrac{\left[(l_1 + l + l_3 + l_4 + l_5) - \frac{l_v}{2} \right] \cdot l_v}{2 E_Y I_v}$. Wherein,

l_v is the length of the isolation beam; E_Y, I_v are the elastic modulus and moment of inertia of the isolation beam, respectively.

D. Theoretical Analysis of Overall Structure

Set the coordinate system along the acceleration direction with the link 1 point as the origin, the deformation of the overall structure is illustrated in Fig. 8. It can be summarized that the resultant displacement of the endpoint of the mass

includes three parts, that is, the isolation beam, support base(including the link 2) and the mass.
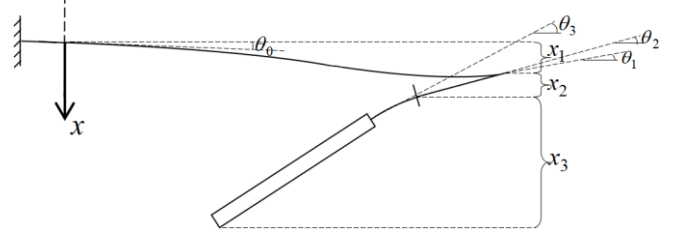
Fig. 8. Diagram of the deformation of overall structure

In accordance with the geometric relationship, the overall coordinate system is actually formed by rotating the coordinates mentioned in the sensing element by the angle θ_0. Thus the actual obliquity of the isolation beam endpoint in the overall coordinate system is derived by the equation (22):

$$\theta_1 = |\theta_v| - |\theta_0| = (\delta_{v\theta} + \delta_{\theta 0}) \cdot F_e = \delta_{\theta 1} \cdot F_e \qquad (22)$$

Meanwhile, according to the coordinate transformation, the displacement of it should be rewritten by the equation (23):

$$x_1 = \delta_{x1} \cdot F_e \qquad (23)$$

where $\delta_{x1} = \delta_{vx} - l_v \cdot \delta_{\theta 0}$

Further, the obliquity and the displacement of the support base is respectively derived as the equations (24) and (25):

$$\theta_2 = \delta_{\theta 2} \cdot F_e \qquad (24)$$
$$x_2 = \delta_{x2} \cdot F_e \qquad (25)$$

where $\delta_{\theta 2} = (\delta_{\theta 1} + \delta_\varphi)$, and $\delta_{x2} = (l_3 + l_4 + 2 l_5) \cdot \delta_{\theta 2}$. Here, it can be approximated that the obliquity of the support base part is equal to the obliquity of the central area of the beam, that is, the actual obliquity of the isolation beam endpoint plus the torsion angle of the transverse 2.

Furthermore, the actual obliquity and the displacement of the mass are respectively derived as the equations (26) and (27):

$$\theta_3 = \delta_{\theta 3} \cdot F_e \qquad (26)$$
$$x_3 = \delta_{x3} \cdot F_e \qquad (27)$$

where $\delta_{\theta 3} = \eta + \delta_{\theta 2}$, and $\delta_{x3} = [\delta_X + (a + l) \cdot \delta_{\theta 2}]$, wherein $X = (1 + a\xi) \cdot x = (1 + a\xi) \cdot \delta \cdot F_e = \delta_X \cdot F_e$, a is the length of mass.

In summary, the resultant displacement of the mass endpoint is derived as the equation (28):

$$x_r = x_1 + x_2 + x_3 = (\delta_{x1} + \delta_{x2} + \delta_{x3}) \cdot F_e = \delta_r \cdot F_e \qquad (28)$$

Therefore, the k_r, which is equal to $1/\delta_r$, is exactly calculated stiffness of the overall mechanical structure of the QVBA.

From the result, it can be demonstrated that the overall structure can be regarded as the series connection of three equivalent springs of the mounting frame, the support base and the sensing element, and the stiffness of overall structure is the equivalent stiffness of the three springs in series. Consequently, the system can be rearranged and simplified to the simplest model, that is the capacitor-mass-spring model shown in Fig. 3.

In addition, the electrostatic force can also be rewritten by the equation (29):

$$F_e = \frac{\varepsilon S U^2}{2(d_0 - x_r)(d_0 - x_r + a \cdot \sin\theta_3)}$$

$$= \frac{\varepsilon S U^2}{2a\xi_r d_0}\left(\frac{1}{d_0 - x_r} - \frac{1}{\frac{d_0}{1-a\xi_r} - x_r}\right) \quad (29)$$

where d_0 is the initial spacing; S is the plate area; and $\xi_r = \delta_{\theta3}/\delta_r$, $\xi_r = \delta_{\theta3}/\delta_r$. Moreover, the equivalent electric stiffness is derived as the equation (30):

$$k_e = \frac{\partial F_e}{\partial x_r} = \frac{\varepsilon S U^2}{2a\xi_r d_0}\left[\frac{1}{(d_0 - x_r)^2} - \frac{1}{\left(\frac{d_0}{1-a\xi_r} - x_r\right)^2}\right] \quad (30)$$

Eventually, in accordance with the analytical methodology of pull-in effect , the equilibrium equation(3) is specified by the following equations (31):

$$\begin{cases} k_r \cdot x_r = \dfrac{\varepsilon S U^2}{2(d_0 - x_r)[d_0 - (1-a\xi_r)\cdot x_r]} \\[2mm] k_r = \dfrac{\varepsilon S U^2}{2a\xi_r d_0}\left[\dfrac{1}{(d_0 - x_r)^2} - \dfrac{1}{\left(\dfrac{d_0}{1-a\xi_r} - x_r\right)^2}\right] \end{cases} \quad (31)$$

By solving the equations, the pull-in displacement is derived as the equation (32):

$$x_{rPL} = \lambda_r \cdot d_0 \quad (32)$$

where $\lambda_r = \frac{2 - a\xi_r - \sqrt{1 - a\xi_r + a^2\xi_r^2}}{3(1-a\xi_r)}$, and the pull-in voltage:

$$U_{rPL} = \gamma_r \cdot d_0^{\frac{3}{2}} \quad (33)$$

where $\gamma_r = \sqrt{\frac{2k_r}{\varepsilon S}\lambda_r \cdot (1-\lambda_r)\cdot[1-(1-a\xi_r)\lambda_r]}$. $\gamma_r = \sqrt{\frac{2k_r}{\varepsilon S}\lambda_r \cdot (1-\lambda_r)\cdot[1-(1-a\xi_r)\lambda_r]}$

III. MODEL VALIDATION AND DISCUSSIONS

A. Simulation and Result

The relative parameters of designed QVBA is summarized in Table 1. as followed. And the acceleration detection range is ±50g.

Taking advantage of the finite element and Matlab, build up a structure-electrostatic coupling field (ignore fringing effect) and substitute parameter values showed in Table 1. The Fig. 9 shows the program of simulation, and Fig. 10 shows the simulation result in the self-test state.

TABLE I.   KEY PARAMETERS OF DESIGNED QVBA

parameter	value	parameter	value
l_1	1.473×10^{-3} m	l_h	1.9×10^{-3} m
l_2	0.13×10^{-3} m	l_v	6.15×10^{-3} m
l	1.0×10^{-3} m	a	2.9×10^{-3} m
l_b	3.0×10^{-3} m	S	9.27×10^{-6} m²
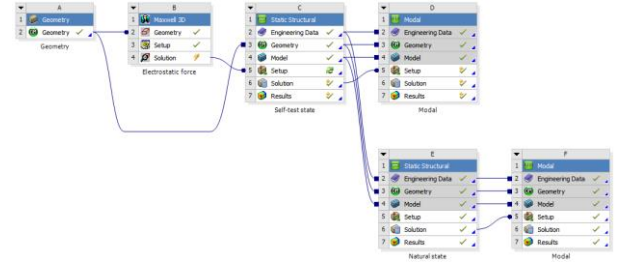
Fig. 9. Simulation program of the model validation
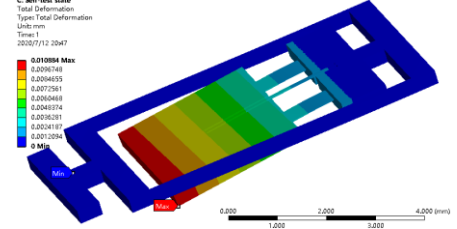
Fig. 10. Simulation result in the self-test state

Because the vibration of the   beam in the width direction is mainly observed during the acceleration measurement, the frequency of the vibration beam in the width direction is used as a comparison standard for specific modal analysis.

According to the result of simulation, Fig. 11 gives the key parameters of the critical pull-in state, that is, the pull-in displacement and the pull-in voltage under different initial spacing. While Fig. 12 shows the vibration frequency of vibration beam under different critical pull-in state with various initial spacing, as well as the corresponding acceleration under natural state.

B. Discussion

The theoretical and simulation results of the relationship between the pull-in displacement and the initial spacing are shown in Fig. 13.

Similarly, Fig. 14 shows the theoretical and simulation results of the relationship between the pull-in voltage and the initial spacing. According to the Fig. 13 and Fig. 14 above, as for pull-in displacement, the error of the theoretical value relative to the simulated value is 0.07%, which is almost negligible; as for pull-in voltage, the error of the theoretical value relative to the simulated value is 1.4%, and the theoretical value is slightly higher than the simulated value. Consequently, the simulation result verifies the correctness of the proposed model.

As for the error of the theoretical pull-in voltage, considering that the theoretical value of pull-in displacement is almost equal to the simulated result and referring to the expression of pull-in voltage and pull-in displacement, the error may directly result from the imperfection of the mechanical stiffness k_r. Through calculation, the theoretical and simulated value of k_r are derived as 139 N/m and 135N/m, respectively. Theoretical value is larger than simulated value. This shows that the assumption is demonstrated. Hence the theoretic modal needs to be developed further.
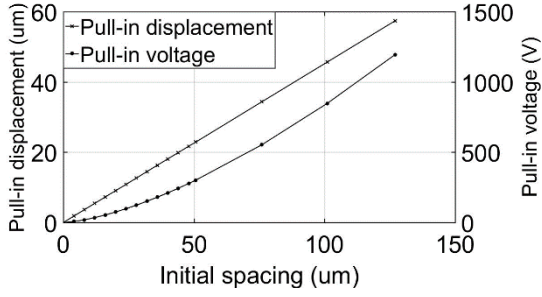
Fig. 11. Key parameters of critical pull-in state with various initial spacing
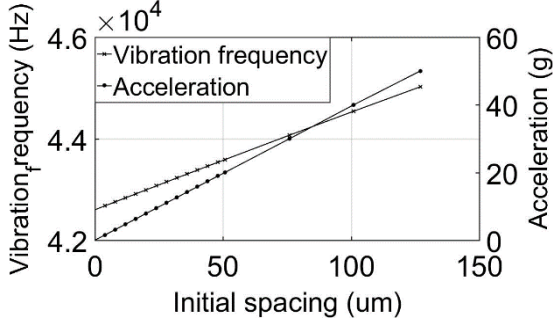
Fig. 12. Vibration frequency under different critical pull-in state and the corresponding acceleration
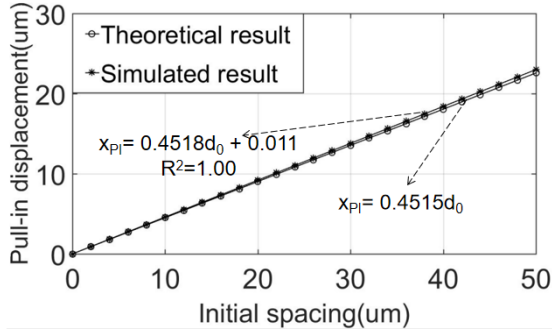
Fig. 13. Relationship between pull-in displacement and the initial spacing
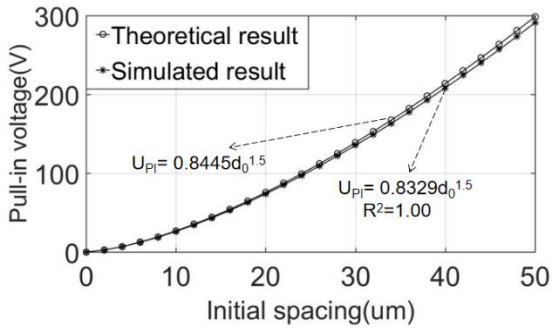
Fig. 14. Relationship between pull-in voltage and the initial spacing

IV. SELF-TEST DESIGN

Fig. 12 shows that the overall displacement is $57.39\mu m$ when acceleration reaches to 50g. To ensure that there is no interference between self-test state and natural state, the initial spacing of electrodes should be wider than the maximum displacement under natural state, $57.39\mu m$.

Here we set the initial spacing as $60\mu m$, and the corresponding pull-in voltage is 387V, equivalent to an applied acceleration of 24g. When the electrode is applied a voltage of 220V, the maximum acceleration simulated is 4.0g.

approximately. Thus, the acceleration detection range is optimized to be $\pm 4g$.

V. CONCLUSION

The overall structure of QVBA and its stress and deformation mechanism are specifically studied in details. The pull-in phenomenon is considered, and an effective theoretical modal of pull-in effect for QVBA is proposed as the equations (32) and (33), and demonstrated to be correct and precise. Wherein, the deformation of isolating element is proved to be non-negligible. A self-test structure is designed and optimized mainly considering the pull-in effect.

REFERENCES

[1] C. Lehrer, L. Frey, S. Petersen, M. Mizutani, M. Takai and H. Ryssel, "Defects and gallium-contamination during focused ion beam micro machining," 2000 International Conference on Ion Implantation Technology Proceedings. Ion Implantation Technology - 2000 (Cat. No.00EX432), Alpbach, 2000, pp. 695-698.

[2] Mengjie Li,Zaifa Zhou,Liyan Yi,Xijie Wang,Saeed Adnan.Design of a test structure based on chevron-shaped thermal actuator for in-situ measurement of the fracture strength of MEMS thin films[J].Nanotechnology and Precision Engineering, 2019,2 (04):163-168.

[3] Panahi, A.; Hossein Sabour, M.; Ghafar-Zadeh, E. A New Non-Invasive Air-Based Actuator for Characterizing and Testing MEMS Devices. Actuators 2020, 9, 24.

[4] R. W. Beegle, R. W. Brocato and R. W. Grant, "IMEMS accelerometer testing-test laboratory development and usage," International Test Conference 1999. Proceedings (IEEE Cat. No.99CH37034), Atlantic City, NJ, USA, 1999, pp. 338-347.

[5] G. F. LaVigne and S. L. Miller, "A performance analysis system for MEMS using automated imaging methods," Proceedings International Test Conference 1998 (IEEE Cat. No.98CH36270), Washington, DC, USA, 1998, pp. 442-447.

[6] Wang Junbo, Shang Yanlong, Chen Deyong, Shi Qiang, Li Guangbei. Closed-Loop Control System of a SOI-MEMS Resonant Accelerometer with Electromagnetic Excitation[J].Nanotechnology and Precision Engineering,2012,10(04): 322-326.

[7] Aikele M, Bauer K , Ficker W , Neubauer F , Prechtel U , Schalk J , Seidal H. Resonant accelerometer with self-test[J]. Sensors and Actuators A (Physical), 2001, 92(1-3), 161-167.

[8] H. C. Nathanson, W. E. Newell, R. A. Wickstrom and J. R. Davis, The resonant gate transistor[J].IEEE Transactions on Electron Devices[J]. 1967,14(3):117-133

[9] Geoffrey Taylor. The Coalescence of Closely Spaced Drops when they are at Different Electric Potentials[J]. 1968, 306(1487):423-434

[10] Batra R C, Porfiri M, Spinello D. Topical review: Review of modelin g electrostatically actuated microelectromechanical systems[J]. Smart Materials & Structures, 2007, 16(6): 23-31

[11] Yang Heng, Bao Minhang, Shen Shaoqun, Li Xinxin, Ren Jianjun. Displacement characteristics analysis of micromechanical structure under the action of electrostatic force(in Chinese) [J]. Journal of Fudan University (Natural Science), 1999, (03): 30-36.

[12] NAYFEH, A.H., YOUNIS, M.I. & ABDEL-RAHMAN, E.M. Reduced-Order Models for MEMS Applications[J]. Nonlinear Dyn 41, 211–236 (2005).

[13] YANG Ting, YANG Gui-yu, LI Qing-feng. Static input-output characteristics for quartz vibrating beam accelerometers[J].Journal of Chinese Inertial Technology,2014,22(03):386-390.

Performance Improvement of Wavelet Noise Reduction Based on New Threshold Function

Shiqi Yu
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:000-0002-5528-1343

Yongrui Qin
Department of Computer Science and Software Engineering
Xi'an Jiaotong-liverpool University
Suzhou, China
ORCID:0000-0002-7783-3399

Jiaxin Gao
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0002-5222-3957

Shiqi Hou
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0003-0366-100X

Fuyong Lyu
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0002-0305-136X

Xuefeng Li*
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0002-5528-1343

Abstract—Acoustic emission (AE) detection, as a non-electrical detection method, is very suitable for effective fault detection of power equipment with a strong electromagnetic field. However, the AE signal collected at industrial sites often contains a lot of interference noise, affecting the analysis and prediction of faults. In this study, a wavelet denoising method based on a new threshold function is proposed, to achieve a noise reduction in the low signal-to-noise ratio (SNR) signals. Simulation experiment results show that the proposed threshold function not only overcomes the shortcomings of the discontinuous hard threshold function, but also solves the constant deviation of the soft threshold function. What's more, the proposed function achieves a good adaptability. When SNR = 10 dB: The SNR of the new threshold is 20.6622, and the RMSE is 0.0026. The SNR of the hard threshold is 20.2246 and the RMSE is 0.0027, compared with the traditional hard threshold method, the SNR of the new threshold is increased by 2.16% and the root mean square error (RMSE) is reduced by 3.7%; the SNR of the soft threshold is 15.5656, and the RMSE is 0.0047, compared with the traditional soft threshold method, the new threshold has a 32.74% increase in SNR and a 40.43% reduction in RMSE. When SNR = -10 dB: The SNR of the new threshold is 4.2602, and the RMSE is 0.0172. The SNR of the hard threshold is 3.8558 and the RMSE is 0.0182, compared with the traditional hard threshold method, the SNR of the new threshold is increased by 10.49% and the RMSE is reduced by 5.49%; the SNR of the soft threshold is 2.1625, and the RMSE is 0.0212, compared with the soft threshold method, SNR is improved by 97% and RMSE is reduced by 18.87%. Performance analyses have proved that the improved wavelet denoising method can obtain a good noise reduction effect. It is very helpful for AE signal analysis with the generally low SNR, which can improve the accuracy of failure identification in subsequent acts.

Keywords—AE signal, wavelet threshold denoising, threshold function, signal-to-noise ratio

I. INTRODUCTION

With its unique advantages as high sensitivity, real-time online, non-destructive, simple operation, rich information, online detection, etc., Acoustic emission (AE) technology can be applied to petrochemical, power, materials, metal processing, civil engineering, transportation, aviation and In the aerospace and other fields [1], certain results have been achieved in monitoring tool damage and wear, evaluating pressure vessel structural integrity, stress state testing of large lifting equipment, research on rotating machinery and aircraft structure monitoring systems, and leak detection. [2]. In practical application, the environmental noise has a large interference and the signal-to-noise ratio (SNR) of the collected AE signal is low, which a great impact on the accuracy of the subsequent analysis results. Therefore, before performing feature extraction, the collected signal is usually subjected to noise reduction processing.

There is a great deal of research that shows the wavelet threshold denoising method can effectively restrain the noise in the process of AE signal acquisition and excellent retained the fault character information in the signals, which has been one of the tools used widely in processing non-stationary signals[3-8]. The main idea is to decompose the AE signal into wavelets of different levels, select an appropriate threshold to process the signal coefficients of each layer, and then reconstruct the signal to achieve the noise reduction effect. Determining the best wavelet function, decomposition layers, threshold, and threshold function are the key points of the wavelet threshold denoising algorithm. Among them, the most widely used threshold noise reduction is the soft and hard threshold noise reduction method proposed by Donoho [9]. However, the soft threshold function suppresses the noise in a constant manner, which achieves the easy removal of effective signals, and the discontinuity of the hard threshold function at the threshold will bring oscillation to the reconstructed signal, both methods have defects in signal noise reduction [10].

In this study, an improved wavelet threshold denoising algorithm is proposed to improve the signal denoising effect. The new approach can solve the typical universal industrial detection problem of strong white noise mixed in the collected AE signals, and provide a higher quality signal for signal characteristic analysis and fault diagnosis.

II. SIGNAL STRUCTURE AND NOISE MODEL

A. AE Signal Structure

The AE signal is a high-frequency signal and exhibits large oscillation characteristics and attenuation characteristics during propagation. The AE signal is a typical attenuation type oscillation pulse signal [11]. A is the signal amplitude, t_0 is the initial moment, f_c is the attenuation oscillation frequency,

---

*Xuefeng Li is the corresponding author. (e-mail: lixuefeng@tongji.edu.cn).

τ is the attenuation time constant, and t is the duration. In general, the AE signals detected on site are mostly oscillation attenuation type signals. Here, the single exponential decay oscillation function y_1 and the double exponential decay oscillation function y_2 are used to simulate the AE signal, and the expression is as follows:

$$y_1 = Ae^{-t/\tau}sin2\pi f_c t \qquad (1)$$

$$y_2 = A(e^{-1.3t/\tau}-e^{-2.2t/\tau})sin2\pi f_c t \qquad (2)$$

B. Interference Noise

Usually, the AE signals detected online contain a lot of interference noise, and these interference noises mainly include narrow-band periodic interference, impulse interference, and white noise [12]. Narrow-band periodic interference is usually denoised by a band-pass filter; impulse interference appears as a short-time pulse signal in the time domain, and as a broadband signal with multiple frequencies in the frequency domain; white noise mainly comes from various circuits Random interference. For the reason that the shielding effect of GIS equipment pipelines and the electromagnetic shielding of detection equipment can be freely designed to reduce the impact of narrow-band periodic interference and pulse interference, this paper mainly studies AE signals under the influence of white noise. Here, the mathematical model of white noise can be expressed as:

$$noise(t) = a * rand(size(t)) \qquad (3)$$

Where a represents the amplitude of white noise and t represents the duration, its power spectral density does not change with frequency, and frequency components are distributed over the entire frequency band. Noise signal such as $s(t)$ can be expressed as:

$$s(t) = x(t) + n(t) \qquad (4)$$

Where $x(t)$ represents a pure signal without noise, and $n(t)$ is a noise signal.

A total of four sets of AE signals are simulated, where signals 1 and 3 satisfy the function y_2, signals 2 and 4 satisfy the function y_1. The amplitudes of the four signals are 1, 0.8, 1.5, 0.3, and the attenuation coefficients are 4 μs, 1 μs, 4 μs, 1 μs respectively. The oscillation frequency is 150 kHz, 200 kHz, 200 kHz, 1 MHz, and the sampling frequency is 1 MHz. Then, the Gaussian white noise is superimposed on the noise-free AE signal, and the Gaussian distribution is satisfied.

III. WAVELET THRESHOLD DENOISING METHOD

The low SNR leads to the weakening of the source characteristics contained in the original AE signals. This study used the wavelet transform method to denoise the noisy AE signal. The denoised signals are used as data samples for feature extraction, training, and recognition, which can improve the accuracy of fault recognition and prediction.

A. Wavelet Transform

The meaning of wavelet transform is that after the function of the mother wavelet is shifted, the inner product of the signal to be analyzed is made at different scales. Thus, for a square integrable real number field $L^2(R)$, the signal to be analyzed $x(t)$, its wavelet transform can be expressed as:

$$W_x(\alpha,\tau) = \frac{1}{\sqrt{\alpha}}\int_{-\infty}^{+\infty} x(t)\psi(\frac{t-\tau}{\alpha})dt \qquad (5)$$

Among them, $\psi(t)$ is the mother wavelet, $\alpha > 0$, called the scale factor, which role is to stretch the basic wavelet $\psi(t)$, τ represents displacement, and $\psi_{\alpha,\tau}(t)$ represents the wavelet sequence. In the wavelet transform, for the signal $x(t)$ to be analyzed, the wavelet sequence obtained after the basic wavelet $\psi(t)$ is scaled and translated is equivalent to the observation window. So the mother wavelet $\psi(t)$ must satisfy:

$$\int_{-\infty}^{+\infty}|\psi(t)|^2 dt < \infty \qquad (6)$$

For the signal $x(t)$ to be analyzed, in $L^2(R)$, the reconstructed expression after decomposition on the wavelet sequence is:

$$x(t) = \frac{1}{C_\psi}\iint_{-\infty}^{+\infty}\frac{1}{a^2}W_x(\alpha,\tau)\psi(\frac{t-\tau}{\alpha})d\alpha d\tau \qquad (7)$$

B. Wavelet Threshold Denoising

The wavelet transform is used to denoise the signal mainly according to the different properties of the wavelet coefficients decomposed in the wavelet domain by the signal and noise, and the appropriate threshold is used to compress or nonlinearly process the wavelet signal containing noise. The processing flow of the wavelet threshold denoising is shown in Fig. 1. The process of denoising a noisy signal is:

1) Wavelet decomposition: According to the principle of wavelet transformation, be fore the wavelet decomposition, a suitable wavelet basis function is used for decomposition. The wavelet coefficients obtained by the decomposition are expressed as: $w_s(j,k) = w_x(j,k) + w_n(j,k)$. Among them, $w_s(j,k), w_x(j,k), w_n(j,k)$ represent the wavelet coefficients of the noisy signal, pure signal, and noise signal on the jth layer of the decomposition level.

2) Wavelet coefficients processing: select the appropriate wavelet threshold and threshold function to process the wavelet coefficients after wavelet decomposition to obtain a new wavelet coefficient $\hat{w}_s(j,k)$.

3) Signal reconstruction: The new wavelet coefficients obtained by the threshold function can be reconstructed by wavelet to obtain the denoised wavelet signal.

C. Noise Removal Effect Evaluation

To judge the evaluation of the denoising effect of noisy AE signals, the following parameters for comparison and analysis:

1) SNR, which is defined as:

$$SNR = 10log\left[\frac{\sum_{i=1}^{n} x^2(t)}{\sum_{i=1}^{n}(x(t)-d(t))^2}\right] \qquad (8)$$

2) The root mean square error (RMSE), which is defined as:
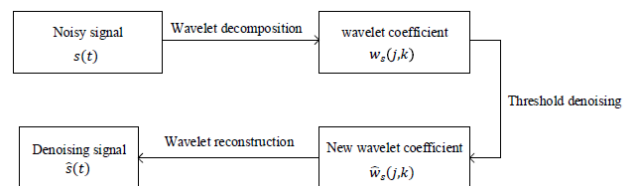
Fig. 1. Flow chart of wavelet threshold denoising.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x(t) - d(t))^2} \qquad (9)$$

Where n represents the number of samples, $x(t)$ represents a pure signal without noise, and $d(t)$ is the output signal after denoising. As defined by the above evaluation indicators, the higher the SNR and the lower the RMSE, the better the denoising effect.

IV. DENOISING METHOD OPTIMIZATION

A. Threshold Function Selection

The commonly used threshold functions for wavelet threshold denoising are divided into hard threshold function and soft threshold function. Equation (10) is a hard threshold function, and (11) is a soft threshold function. The function expression is as follows:

$$\hat{w}_s(j,k) = \begin{cases} w_s(j,k), & |w_s(j,k)| \geq \lambda \\ 0, & |w_s(j,k)| < \lambda \end{cases} \qquad (10)$$

$$\hat{w}_s(j,k) = \begin{cases} sign(w_s(j,k)) * (|w_s(j,k)| - \lambda), & |w_s(j,k)| \geq \lambda \\ 0, & |w_s(j,k)| < \lambda \end{cases} \qquad (11)$$

Where $w_s(j,k)$ is the wavelet coefficients decomposed by the original signal, λ is the set threshold, and $\hat{w}_s(j,k)$ is the wavelet coefficient after threshold processing.

For the hard threshold, the function has a discontinuity while $w_s(j,k) = \lambda$. When the signal is reconstructed later, a larger oscillation will occur. Compared with the hard threshold function, the soft threshold function is continuous in the definition domain. However, when the absolute value of $w_s(j,k)$ is greater than λ, there will be a constant deviation between $\hat{w}_s(j,k)$ and $w_s(j,k)$, which will affect the accuracy of the reconstructed signal.

To ensure the continuity of the function and improve the constant deviation between $\hat{w}_s(j,k)$ and $w_s(j,k)$, based on the expressions of the hard threshold function and the soft threshold function, this paper uses the concavity and convexity of the exponential function to reconstruct a new The threshold function of the expression is as follows:

$$\hat{w}_s(j,k) = \begin{cases} w_s(j,k) - \frac{0.5\lambda^{m+1}}{w_s(j,k)^m}, & w_s(j,k) > \lambda \\ \frac{0.5sign(w_s(j,k))\left(sign(w_s(j,k)) * w_s(j,k)\right)^{m+1}}{\lambda^m}, & |w_s(j,k)| \leq \lambda \\ w_s(j,k) + \frac{0.5\lambda^{m+1}}{(-w_s(j,k))^m}, & w_s(j,k) < -\lambda \end{cases} \qquad (12)$$

Where m (m$\geqslant$1) is a positive integer, which is the adjustment coefficient of the exponential function. For signals with different SNRs, the adjustment coefficient can be used to obtain the threshold function with the best denoising effect. When $\lambda = 4$, m respectively takes 2, 4, 6, 8, and 10 are obtained as shown in Fig. 2. The new function domain defined is the entire real number domain, which is the same as the traditional soft threshold function and the hard threshold function. At $w_s(j,k) = \lambda$, because $\hat{w}_s(j,k)_- = \hat{w}_s(j,k)_+$, so the new function is continuous at $w_s(j,k) = \lambda$, which can improve the oscillation of the traditional hard threshold function during signal reconstruction due to discontinuity. When$|w_s(j,k)| \geq$

λ, $|\hat{w}_s(j,k) - w_s(j,k)| = 0.5 * \lambda^{m+1}/w_s(j,k)^m$, if m is constant, follow as $w_s(j,k)$ increases, the difference gradually decreases, improving the soft threshold function when $|w_s(j,k)| \geq \lambda$, $\hat{w}_s(j,k)$ and $w_s(j,k)$. The constant deviation existing between them improves the accuracy of signal reconstruction.

B. Threshold Selection

In the wavelet threshold denoising, the threshold is a very important parameter. When the threshold is too much set, part of the effective feature signal is filtered as a noise signal during the signal processing; if the threshold is set too small, a large amount of decomposed wavelet coefficients remain. The noise signal will cause signal distortion. Therefore, the selection of thresholds has a very important effect on the denoising quality of the wavelet threshold denoising method.

In 1994, Donoho and Johnstone proposed a general formula for setting thresholds[9], expressed as:

$$\lambda = \sigma\sqrt{2lnN} \qquad (13)$$

Where σ represents the noise variance and N represents the length or size of the signal. However, in practice, the standard deviation of the noise is unknown. Therefore, the threshold is usually selected by the estimation method, and the commonly used estimation method is shown in (14):

$$\sigma = \frac{median(|w_{j,k}|)}{0.6745} \qquad (14)$$
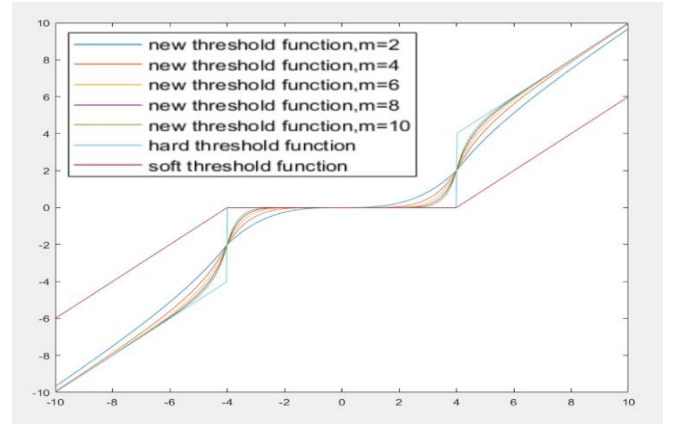
(a)
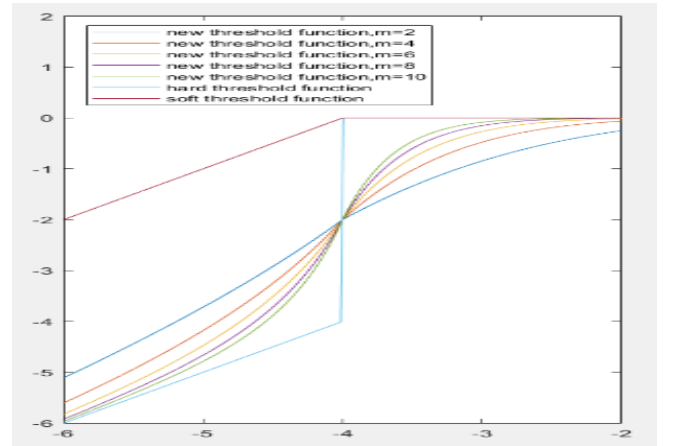
Fig. 2. Threshold functions for wavelet threshold denoising. (a)Soft, hard, and new threshold function (m=2, 4, 6, 8, 10), (b) Enlarged view of local details.

Where $w_{j,k}$ represents the decomposition coefficient at each scale. Equation (13) is related to the length of the signal. When the length N of the signal is large, the threshold also obtains a large value. The wavelet coefficients obtained after denoising will be mostly set to zero, and the wavelet denoising is degraded to a low-pass filter at this time.

Although the general threshold formula has good theoretical support, there are still certain defects in the denoising process. For the coefficients decomposed by each layer, the threshold is set to a constant value, which lacks traceability. Therefore, to better deal with wavelet coefficients, the threshold formula as shown in (15) is used:

$$\lambda = \sigma \frac{\sqrt{2 \ln N}}{\ln (j+1)} \qquad (15)$$

Where j represents the decomposition scale so that the threshold is set differently at different decomposition scales. As the decomposition scale increases, the wavelet coefficient of the noise decreases accordingly, which can distinguish the noise from the original signal more effectively.

Use different threshold functions to denoise the noisy signal, and the denoising effect is shown in Fig. 3.

As can be seen from the Fig. 3, the hard threshold function and the new threshold function are better than the soft threshold signal in the restoration of the effective signal, and retain the effective information of the signal to a greater extent.

(a)

(b)

(c)

(d)
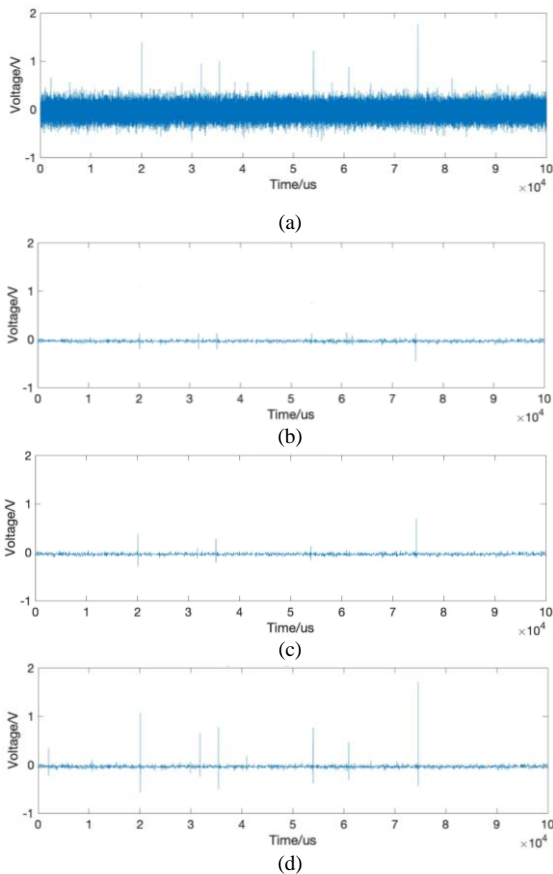
Fig. 3. The polluted signals and the denoised signals using different methods. (a) the polluted signals. (b) the denoised signals using the soft threshold function. (c) the denoised signals using the hard threshold function. (d) the denoised signals using the new threshold function.

V. DENOISING EFFECT ANALYSIS

Based on the method described above, Gaussian white noise with different intensity is added to the noiseless analog AE signal, and its SNR is expressed as a set of [-10 dB, -5 dB, 0 dB, 5 dB, 10 dB]. Then, use the wavelet threshold denoising method to denoise, perform 6-layer wavelet decomposition on the noise signal, and use db2 as the wavelet basis function, respectively choose the hard threshold function of (10), the soft threshold function of (11), the new threshold function of (12) proposed by this study as the wavelet function, and function (15) is selected as the improved threshold for processing the wavelet coefficients of each layer.

Fig. 4 shows the comparison and analysis of the noise reduction effect of three threshold functions, where m is set to 10 in the new threshold function. As shown in Figure 4(a), the noise AE signal from high SNR (10 dB) to low SNR (-10 dB) is denoised here, and the displayed datas are the SNR improvement effect after signal denoising, It can be seen that the three threshold functions have a good denoising effect on low SNR signal, and the new threshold function has the best denoising effect. From the data of high SNR signal, the noise reduction effect of the new threshold function and the hard threshold function is good, and the noise reduction effect of the soft threshold function is not ideal. Figure 4(b) shows the comparison of the RMSE parameters after signal noise reduction. The new threshold function has the best noise reduction effect, while the hard threshold function and soft threshold function have different noise reduction effects in

(a)

(b)
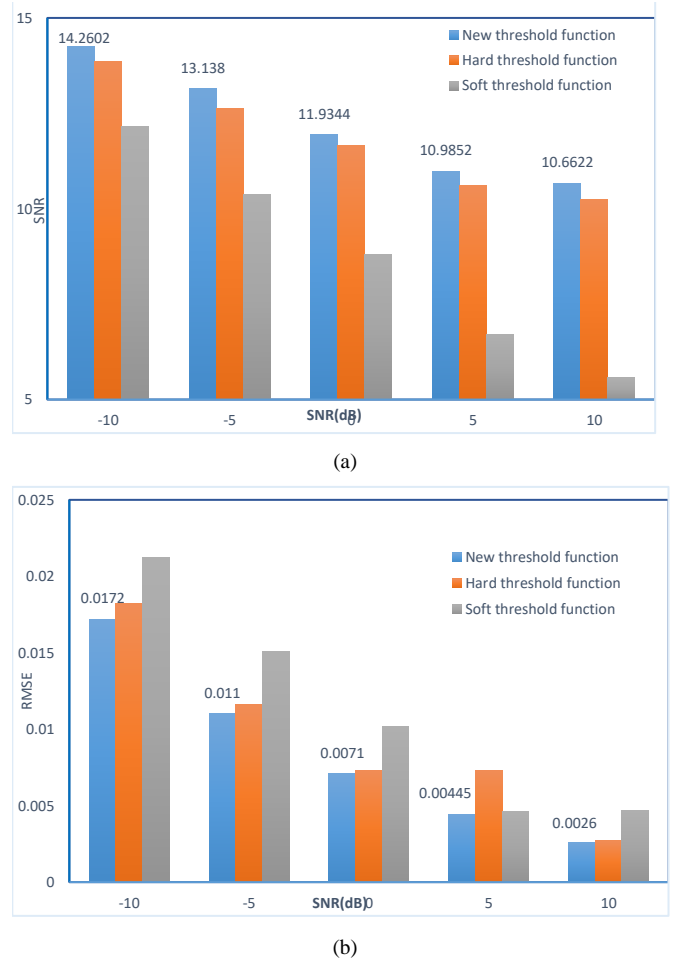
Fig. 4. Denoising the effect of three different threshold functions. (a) SNR comparison, (b) RMSE comparison.

different noisy AE signals. In summary, through a comparative analysis, the new threshold function shows a better ability to reduce noise and retain the original AE signal than the traditional threshold function.

VI. Conclusion

The phenomenon of AE has been used as a powerful tool for detecting, locating, or evaluating the damage in various applications. In this study, a wavelet denoising method with an improved threshold function is adopted to denoise the AE signal. The denoising effect of soft threshold function, hard threshold function, and new threshold function on noisy AE signal under the background of different SNR is analyzed. Research has shown that the proposed new threshold function has better denoising effect and signal feature retention ability, and can be well applied to the denoising of non-stationary signals such as AE signals. For the typical problem that the AE signal collected in the industrial field is mixed with a lot of noise, it is of great significance and value to provide a higher quality signal for fault diagnosis.

References

[1] T. Sheng, H. Wei, H. Zhinan and M. Tianqi, "Application of wavelet analysis in acoustic emission signal processing of coal and rock fracturing," Engineering and Technology, vol. 12, pp. 190–191, June 2017.

[2] Y. Li, "Acoustic emission monitoring of main intake duct of large structure," in 2009 12th Chinese Acoustic Emissions Seminar in China,Nanjing, 2009.

[3] W. MT, "Wavelet transform based on Meyer algorithm for image edge and blocking artifact reduction," Information Sciences, vol. 474, pp. 125-135, 2019.

[4] R. Q. Quiroga, "Obtaining single stimulus evoked potentials with wavelet denoising," Physica D: Nonlinear Phenomena, vol. 145, pp. 278-292, November 2000.

[5] D. Giaouris, J. W. Finch, O. C. Ferreira, R. Kennel, G. Elmurr, "Wavelet Denoising for Electric Drives," IEEE Transactions on Industrial Electronics, vol. 55(2), pp. 543–550, 2018.

[6] X. Liu, Z. Liu, X. Li, M. Rao, L. Dong, "Wavelet threshold de-noising of rock acoustic emission signals subjected to dynamic loads," Journal of Geophysics and Engineering, vol. 15(4), pp. 1160–1170, 2018.

[7] S. Kumar, B. Singh, "Quantification of tool chatter and metal removal rate using wavelet denoising and statistical approach," Noise & Vibration Worldwide, vol. 49(2), pp. 62-81, 2018

[8] J. Liu, F. L. Qin,"Improved wavelet threshold method and its application Application of seismic data noise reduction," Geophysical and Geochemical Exploration, vol. 44(4), pp. 1-6, 2020

[9] D. L. Donoho, I. M. Johnstone, "Ideal spatial adaption via wavelet shrinkage," Biometrika, vol. 81, pp. 425–455, 1994.

[10] Z. P. Liu, L. Zhang, J. Carrasco, "Vibration analysis for large-scale wind turbine blade bearing fault detection with an empirical wave-let thresholding method," Renewable Energy, pp. 146, 2020.

[11] D. Dengwei, T. Cheng, G. Wensheng, L. Weidong, Y. Senjing, Z. Yuming, "Spectrum characteristics and propagation characteristics of typical partial discharge in GIS," High voltage technology, vol. 40(10), pp. 3243-3251, 2014.

[12] L. Qi, L. Wei, "Simulation Analysis of PD Signal of Partial Discharge in Power Transformer," Heilongjiang Power, vol. 2017(2) , pp. 145-148, 2017.

[13] B. Jicheng, "Application of Median Filtering and Wavelet Transform in Signal Denoising," Application of Median Filtering and Wavelet Transform in Signal Denoising, vol. 2012(28), pp. 294-295, 2012.

Optimal Design for New Support Structure of Quartz Resonators

Jing Ji *
School of Electro-Mechanical Engineering
Xidian University
Xi'an, China
jingji@xidian.edu.cn

Meng Zhao
School of Electro-Mechanical Engineering
Xidian University
Xi'an, China
mzhao@xidian.edu.cn

Toshitsugu Ueda
IPS Research Center
Waseda University
Fukuoka, Japan
tueda@waseda.jp

Abstract—**Frequency generators and sensors increasingly require quartz resonators vibrating in high-frequency with high Q-factor. Currently, most of high-frequency resonators vibrating in above 150 MHz are overtone mode type, which is not so easy to achieve high Q-factor comparing with fundamental-type resonators. Since the quartz etching process technology has been improved, it becomes possible to fabricate thin plate of quartz resonator to achieve a high frequency value of fundamental mode. The fundamental mode resonators are gradually applied. To achieve a high Q-factor value, the dimension should be optimized to restrain vibration energy under the electrodes of resonators. Many efficient achievements have been acquired in this research field. However, many works focus on electrode dimension, and do not describe much for support part. In this work, we applied energy trapping analysis to optimizing support structure for a fundamental type quartz resonator with high frequency value. A new structure with symmetric support parts was designed and optimal dimensions were presented.**

Keywords—quartz resonator, optimal design, support, energy trapping

I. INTRODUCTION

High-frequency AT-cut quartz resonators with high Q-factor are increasingly required by frequency generating and frequency controlling devices. Currently, most of high-frequency resonators vibrating in above 150 MHz are overtone mode type[1], which is not so easy to achieve high Q-factor comparing with fundamental-type resonators. Since the quartz etching process technology has been improved, it becomes possible to fabricate thin plate of quartz resonator to achieve a high frequency value of fundamental mode. The fundamental mode resonators are gradually applied.

To achieve a high Q-factor value, the dimension should be optimized to restrain vibration energy under the electrodes of resonators[2]. Many efficient achievements have been acquired in this research field. However, many related works focus on electrode dimension, and do not describe much for support part[3].

In our former work, we first applied energy trapping analysis to optimizing single support part structure for a fundamental-type quartz resonator [4]. In this paper, a new structure with two asymmetrical support parts was firstly designed and optimal dimensions were presented.

II. RESONATOR

The structure of our resonator is illustrated in Fig. 1. The quartz substrate can be introduced in two parts: Rectangular solid vibration part (1440×1500×10 µm) and fixed prism support parts. The support parts are prisms with trapezoidal

*Jing Ji is the corresponding author. (e-mail: jingji@xidian.edu.cn).

sections. The prisms heights L and R are the varying parameters to be optimized. The trapezoid height is 200 µm. The lengths of the parallel sides of trapezoid equal 10 and 100 µm respectively.

The vibration part is covered partially by a rectangular upper exciting electrode (1000×1000×0.13 µm) and covered fully by a rectangular lower grounded electrode (1440×1500×0.13 µm). The electrodes consist of Cr and Au layers (Cr 0.03 µm, Au 0.1 µm).
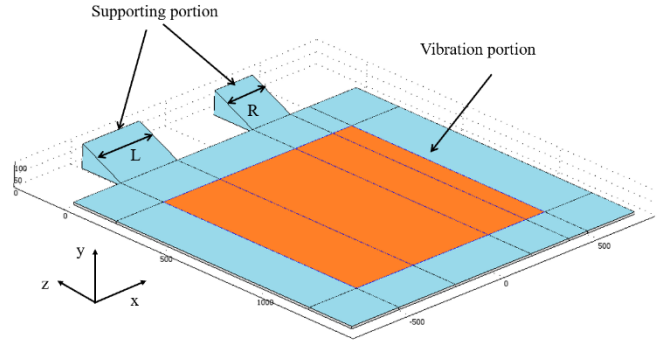
Fig. 1: resonator structure, where L and R is the parameters of support parts to be optimized.

This resonator works in a fundamental thickness-shear (TS) vibration that is most common vibration mode for quartz resonators with high-frequency value. The theoretical TS vibration frequency value [5] and the thickness of quartz plate are inversely proportion and defined as:

$$f_0 = \frac{\sqrt{G/\rho}}{2b} \qquad (1)$$

where ρ denotes quartz density, G denotes the shear modulus ($\rho = 2649 kg/m^3$, $G = 29.01 GPa$), and b denotes the quartz plate thickness ($b = 10\ \mu m$).

The non-optimized resonator has two asymmetrical support parts (L=300 µm, R=200 µm). To effectively confine vibration energy in the central excitation electrode, a new symmetric support parts structure was adopted. The parameters L and R simultaneously varied from 200 to 500 µm.

III. MODEL

We chose linear rectangular elements to establish finite element model to ensure the element quality and calculation accuracy. For the AT-cut quartz plate, parameters are given as follows [5].

Elasticity matrix:

$$C_E = \begin{pmatrix} 86.74 & -8.25 & 27.15 & -3.66 & 0 & 0 \\ -8.25 & 129.77 & -7.42 & 5.7 & 0 & 0 \\ 27.15 & -7.42 & 102.83 & 9.92 & 0 & 0 \\ -3.66 & 5.7 & 9.92 & 38.61 & 0 & 0 \\ 0 & 0 & 0 & 0 & 68.81 & 2.53 \\ 0 & 0 & 0 & 0 & 2.53 & 29.01 \end{pmatrix}$$

$$GPa$$

Dielectric constant matrix:

$$\mathcal{E}_{rS} = \begin{pmatrix} 39.21 & 0 & 0 \\ 0 & 39.82 & 0.86 \\ 0 & 0.86 & 40.42 \end{pmatrix} \times 10^{-12} F/m$$

Piezoelectric constant matrix:

$$e = \begin{pmatrix} 0.171 & -0.152 & -0.0187 & 0.067 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.108 & -0.095 \\ 0 & 0 & 0 & 0 & -0.0761 & 0.067 \end{pmatrix}$$

$$C/m^2$$

The parameters of the electrodes are given in Table I .

TABLE I.   ELECTRODES MATERIAL PARAMETERS

Material	Density (kg/m³)	Young's Modulus (GPa)	Poisson Ratio
Au	19320	77.2	0.44
Cr	7190	270	0.21

Mesh condition are given in Table II.

TABLE II.   MESH CONDITION

Meshes	Quartz	Exciting electrode	Grounded electrode
X-direction	70	44	70
Y-direction	10	1	1
Z-direction	67	50	67

Fig. 2 shows the finite element model used rectangular cuboid type mesh with approximately 235,000 degrees of freedom and the calculated result of main TS vibration mode.
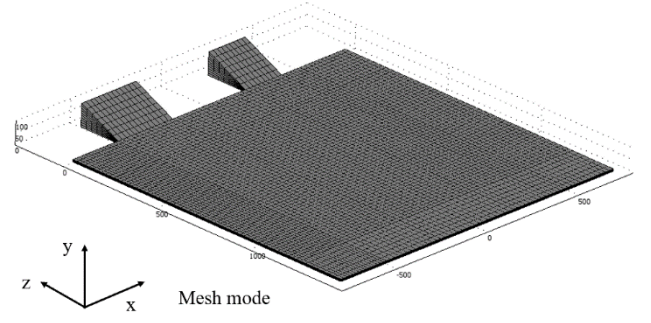
Fig. 2: Finite element modeling.

IV. OPTIMIZATION

When the mass of exciting and grounded electrodes is loaded, the resonance frequency under the exciting electrode is much lower than the value of the region without electrodes. This causes TS vibration mode to be excited under the electrode region and corresponding vibration to decay exponentially outside. This is energy trapping effect [6]. The numerical fundamental TS mode frequency values can be calculated according to the Sauerbrey's equation [7][8] for different electrode conditions:

$$f = f_0 + \Delta f \qquad (2)$$

$$\Delta f = \frac{-2f_0{}^2 \Delta M}{A\sqrt{\rho G}} \qquad (3)$$

where Δf denotes the frequency shift, f_0 denotes the resonance frequency of the resonator without electrode loading, ΔM denotes the mass of electrode, and A denotes the surface area of electrode.

In order to achieve a high Q-factor, the dimension of support parts should be optimized so that vibration energy can be confined efficiently under the excitating electrode.

The x-displacement surface distribution in main mode was calculated for the non-optimized resonator and the new designed resonators. Fig. 3 shows, for the non-optimized resonator, vibration energy was not efficiently confined in the central excitation electrode region but concentrated on one end. For resonators with symmetric structure, energy trapping was improved (see Fig. 4, 5, 6, and 7).

In order to quantitatively analyze the energy trapping effect, we introduced a parameter called energy trapping rate:

$$R = \frac{E_e}{E_e + E_o} \qquad (4)$$

Where E_e is the vibration energy under exciting electrode , and E_o is the vibration energy outside.
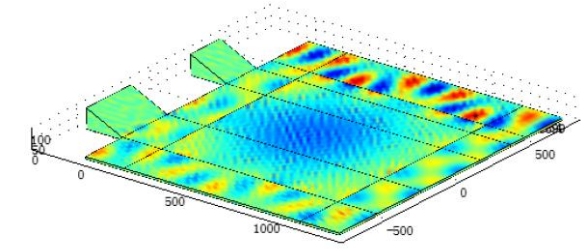
Fig. 3: x-displacement distribution in main mode (L=300 µm, R=200 µm).
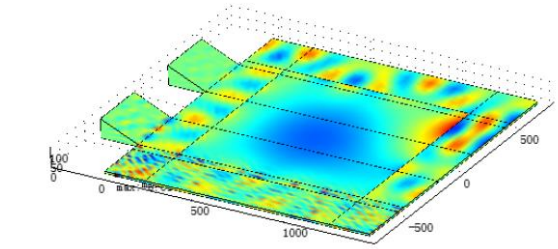
Fig. 5: x-displacement distribution in main mode (L=300 µm, R=300 µm).
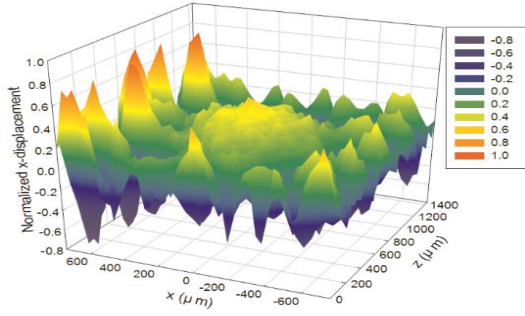
Fig. 4: x-displacement surface distribution in main mode (L=200 µm, R=200 µm).
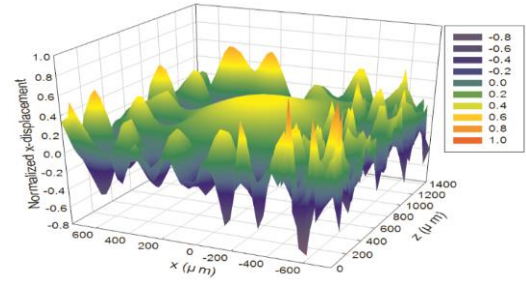
Fig. 6: x-displacement distribution in main mode (L=400 µm, R=400 µm).
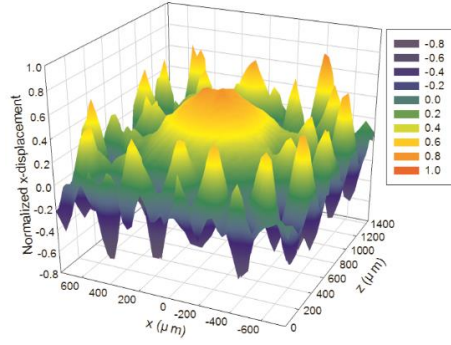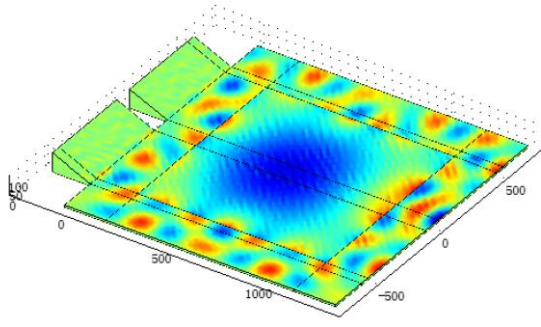
Fig. 7: x-displacement distribution in main mode (L=500 µm, R=500 µm).

According to Table. III, relatively high energy trapping rates were obtained for new designed resonators No.3 and No.4. Thus, optimal sizes of support part were given as follows, L=R=400 µm and L=R=500 µm.

TABLE III. ENERGY TRAPPING PERFORMANCE.

	Former design	**New design**			
		No.1	*No.2*	*No.3*	*No.4*
L (µm)	300	200	300	400	500
R (µm)	200	200	300	400	500
Energy Trapping Rate (%)	42.98	46.48	44.36	64.35	54.37

V. CONCLUSIONS

In this work, we optimized a new designed support structure for high-frequency fundamental-type AT-cut quartz resonators. Optimal dimensions of support part were determined by improving energy trapping effect, which has relatively high energy trapping rates. The validation of our work was ensured by improvement of the optimized resonator's performance.

Ideally, quartz resonator works in a pure fundamental TS vibration. However, quartz plate boundary caused the fundamental vibration mode always couples with some spurious vibration modes [9]. Therefore, we will focus on reducing vibration couplings in future, and fabricating resonators with varing dimensions and measured their Q-factors to further verify the optimal work.

In our future works, we will also consider how to how to improve the temperature-frequency coefficient performance, and how to optimize quartz resonators with different cut angle

REFERENCES

[1] Yong, Yook-Kong. (2012). Resonator Q increase and noise reduction in third overtone thickness shear resonators. 2012 IEEE International Frequency Control Symposium, IFCS 2012, Proceedings. 1-6. 10.1109/FCS.2012.6243721.

[2] Abe, Takashi & Kishi, Hiroki. (2011). A Gaussian-shaped AT-cut quartz crystal resonator. Sensors and Actuators A Physical. 166. 173-176. 10.1016/j.sna.2010.04.005.

[3] Lu, Feng & Lee, Hp & Lu, Pin & Lim, S.P.. (2005). Finite element analysis of interference for the laterally coupled quartz crystal microbalances. Sensors and Actuators A: Physical. 119. 90-99. 10.1016/j.sna.2004.09.013.

[4] Jing Ji, Meng Zhao, Hiroshi Oigawa, & Toshitsugu Ueda. (2018). Optimal Design of Support for Planoconvex Quartz Resonator, Sensors and Materials, 30. 5. 063–1071. 10.18494/SAM.2018.1832

[5] Wang, Ji & Li-jun, Shen. (2005). Exact thickness-shear resonance frequency of electroded piezoelectric crystal plates. Journal of Zhejiang University SCIENCE A. 6. 10.1007/BF02857345.

[6] Yamada, Ken & Seto, Shuichi. (2010). Use of Trapped-Energy Mode of Backward-Wave-Type Thickness Vibration for Liquid-Level Sensing. Japanese Journal of Applied Physics. 49. 1439-1441. 10.1109/ULTSYM.2010.5935607.

[7] Yang, Jiashi & Zhou, Honggang & Zhang, Weiping. (2005). Thickness-shear vibration of rotated Y-cut quartz plates with relatively thick electrodes of unequal thickness. IEEE transactions on ultrasonics, ferroelectrics, and frequency control. 52. 918-22. 10.1109/TUFFC.2005.1503978.

[8] Sauerbrey, G. Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung. Z. Physik 155, 206–222 (1959). https://doi.org/10.1007/BF01337937

[9] Wang, Ji & Pan, Qiaoqiao & Yang, Lijun & Chao, Min-Chiang. (2010). Design of quartz crystal resonators with an analytical procedure based on the Mindlin plate theory. Proceedings - IEEE Ultrasonics Symposium. 1246 - 1249. 10.1109/ULTSYM.2010.5935851.

Behavior Anomaly Detection Fused with Features of Mel Frequency Cepstrum Coefficients

Fuyong Lyu
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0002-0305-136X

Zhouhang Yang
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0002-1947-0091

Lujie Wang
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0001-7749-9725

Qi Zhou
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0001-9076-0094

Shiqi Hou
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0003-0366-100X

Xuefeng Li*
College of Electronic and Information Engineering
Tongji University
Shanghai, China
ORCID:0000-0002-5528-1343

Abstract—**In recent years, with the rise of artificial intelligence, the research in the field of pattern recognition is developing rapidly. Its supporting algorithms mainly include neural network and support vector machine (SVM), which are widely used in various fields such as industrial detection, aviation, medical treatment, finance, internet and so on. On the one hand, the effect of pattern recognition is related to the selection and optimization of the algorithm, on the other hand, it is closely related to the noise reduction methods to improve the signal quality, the feature extraction and feature selection in feature engineering. Mel frequency cepstrum coefficient (MFCC) is widely used in the field of speech recognition, and has achieved good recognition effect for low-frequency speech signal pattern recognition, but its pattern recognition in ultrasonic field is rarely studied. In this paper, a method of extracting MFCC in ultrasonic frequency range and applying it to SVM pattern recognition and fault prediction is proposed. In the experimental verification stage, GIS-PD model is used to collect data of acoustic emission (AE) signal induced by partial discharge (PD) of gas insulated switchgear (GIS), and MFCC eigenvalues are extracted in the ultrasonic frequency range of 20-200 kHz. Finally, MFCC and optimized selected time-domain and frequency-domain eigenvalues are used for SVM pattern recognition to monitor and warn the running state of GIS equipment. Based on the comprehensive scheme proposed in this study, the overall accuracy rate is improved to ninety percent after fusion of MFCC eigenvalues, which is improved by 22.5 percent compared with only using time-frequency domain eigenvalues. At present, the research on MFCC extraction in the ultrasonic frequency range of 20-200 kHz and its application in SVM pattern recognition has not been reported. This study will promote the research of MFCC in the field of ultrasonic detection, and provide reference for the research in related fields.**

Keywords—*MFCC, optical fiber sensor, partial discharge, pattern recognition, SVM*

I. INTRODUCTION

High voltage electrical equipment is the core part of the power system. The online monitoring and fault diagnosis are very important to ensure the operation safety of the power system. Gas insulated switchgear (GIS) has the advantages of compact structure, strong versatility of components, convenient installation and high reliability, which can improve the operation stability of power supply system and meet the increasingly stringent requirements of modern society for power industry and power supply system security. Therefore, it is widely used in the field of high voltage and ultra-high voltage [1-3]. However, reduction of safety distance between internal components due to compact equipment structure, continuous improvement of working voltage level, and insulation defects in manufacturing, transportation, construction and installation of GIS equipment will cause partial discharge (PD) in GIS internal equipment, which will lead to deterioration of GIS equipment performance. Serious PD will even lead to breakdown of insulation equipment, which directly threatens the safe and stable operation of power system [4]. In particular, GIS equipment is mostly used in high-voltage substation and high-voltage transmission trunk line, which undertakes the task of regional to global transmission and distribution. Once the equipment fails, the influence range of power failure accident will expand rapidly. Therefore, it is urgent to realize the early detection and maintenance of GIS insulation defects through PD signal detection and defect type identification, and to develop an intelligent fault diagnosis system capable of on-line real-time detection of high-voltage power equipment faults, which has become a hot topic in the field of power industry [5-7].

Acoustic emission (AE) detection, as a non-electrical detection method, is very suitable for effective fault detection of power equipment with strong electromagnetic working environment. However, in the field of GIS equipment fault detection, due to its complexity and airtight structure, it is still a challenge to study the pattern recognition of PD fault source induced by insulation defect [8-10]. Scholars have used many methods to conduct data mining and pattern recognition for PD fault of GIS equipment to realize fault information analysis and diagnosis [11-12]. Among them, support vector machine (SVM) is a new machine learning method developed on the basis of statistical learning theory, which has obvious training effect for small sample data and high recognition accuracy, and is increasingly applied to pattern recognition of PD fault in GIS equipment [13]. On the other hand, eigenvalue is an important factor that affects the recognition accuracy, so how to select the features that are strongly related to the signal source is very important. Mel frequency cepstrum coefficient (MFCC) is widely used in speech recognition field because it can reflect the auditory characteristics of human ear and has a good recognition effect. What's more, it is applied to the

*Xuefeng Li is the corresponding author. (e-mail: lixuefeng@tongji.edu.cn).

industrial field by more and more researchers [14-17]. However, at present, the public research results are based on the low-frequency sound signal for pattern recognition, and the environmental noise or system noise also exists in the same frequency band, which brings great interference to the detection and location of fault sources [18-21].

Therefore, based on the optical fiber AE sensor, this study continuously monitors and analyzes the PD induced AE signal of GIS equipment, extracts MFCC in the ultrasonic frequency range for SVM model training and testing, and then explores the prediction method of potential insulation failure in high-voltage electrical equipment. The research results show that the developed fiber-optic AE sensor, combined with MFCC and the optimized time-domain and frequency-domain eigenvalues, can be used to train SVM to improve the rate of recognition, effectively monitor and forewarn the operation status of GIS equipment, which has positive significance and good application prospects for ensuring the safe operation of power system.

II. EXPERIMENT

The existing operation and maintenance experience of GIS equipment shows that when PD occurs, it is accompanied by a variety of physical and chemical phenomena. The generated high-energy electrons will form ionized gas channels, which will cause local instantaneous heating to produce gas detonation. The current existing in a very short period of time will also cause rapid expansion and contraction of the insulation medium, and its changing process will also generate ultrasonic waves in the form of vibration.

The optical channel of the optical fiber pressure sensor based on polarization modulation used in this study is shown in Fig. 1. The sensing unit of the sensor is constructed by panda polarization maintaining fiber (PMF). The linearly polarized light enters the fiber at an angle of 45° to the slow axis (x axis) of the PMF. Therefore, the light intensity and phase of the slow axis and fast axis (y axis) are the same at the input end of PMF. The output light of the PMF is separated into two vertically polarized lights by the splitter, one of which is rotated by 90° and then coupled with the other. The light intensity at the output end can be expressed as follows:

$$I_o = \frac{1}{2}E_i^2 \cdot (1 + cos(k_0 LB)) \tag{1}$$

Where I_o is the output light intensity, E_i is the incident field strength of two orthogonal polarization states, k_0 is the wave number, L is the length of the PMF, and B is the birefringence of PMF. When PMF is affected by the external force, the photoelastic effect will change its optical performance (mainly effective refractive index), thus
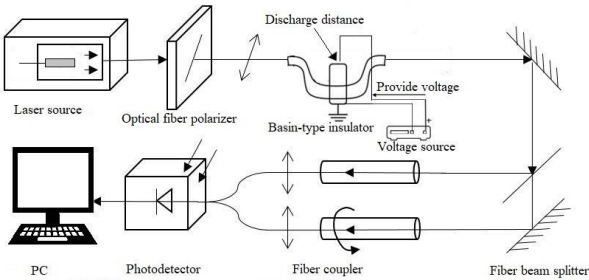
Fig. 1. Optical fiber ultrasonic sensor and its detection system.

changing the output light intensity. By demodulating the intensity signal of the output light, we can obtain the intensity and frequency of the external forces that affect the PMF.
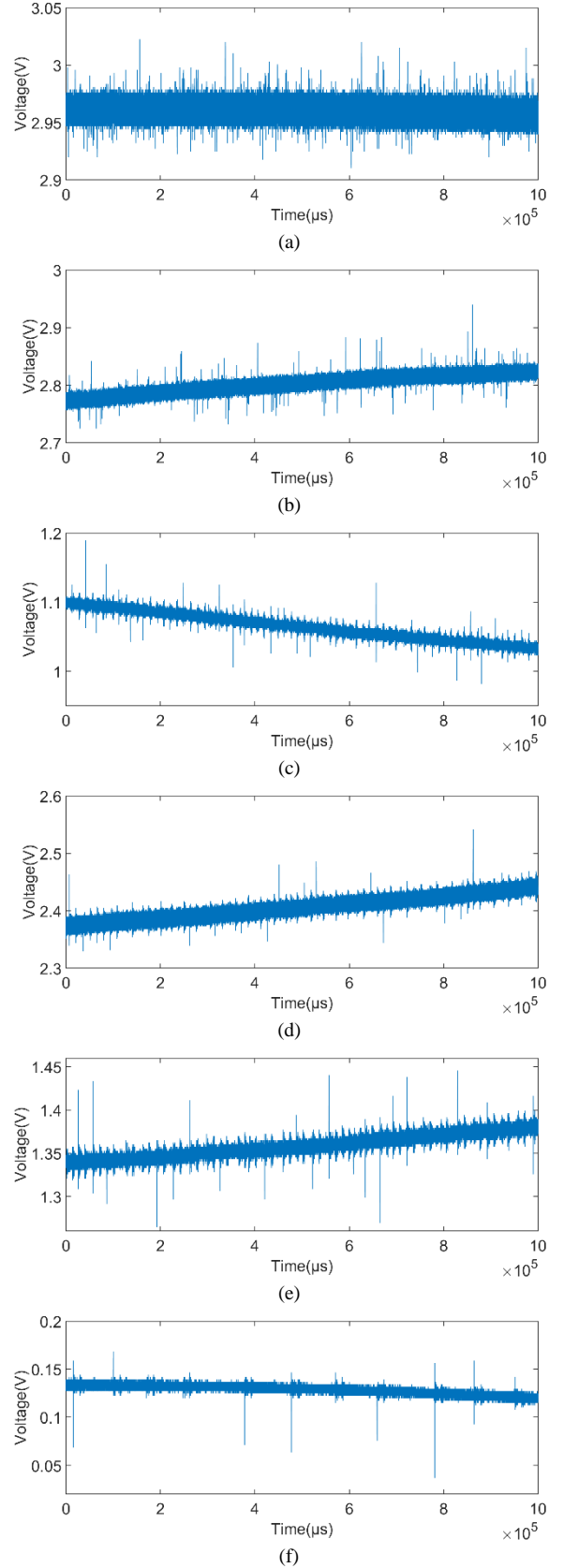
Fig. 2. The collected signal. (a) 1 mm, (b) 2 mm, (c) 3 mm, (d) 4 mm, (e) 5 mm, (f) 6 mm.

90

The established detection system is divided into an optical subsystem and a discharge subsystem. The optical subsystem mainly includes: laser light source, optical fiber polarizer, optical fiber rotator, sensing unit, optical fiber beam splitter, optical fiber coupler, photoelectric detector and data acquisition and processing terminal. The sensing unit uses a 1.0-meter-long PMF (SM15-PS-U25D, Fujikura, Japan), and is sealed in the U-shaped groove around the basin-type insulator with Vaseline as the binder to collect the AE signal induced by PD. The discharge subsystem uses a pin-to-plate discharge model, which is mainly composed of a high-voltage generator, a discharge tip, and an insulator. The maximum discharge voltage of the high-voltage generator (DG1000Z, Rigol, China) is 15 kV. When the tungsten alloy discharge needle as the high-voltage electrode approaches the inner conductor of the basin-type insulator, the discharge channel is formed, which is used to simulate the working condition of the high-voltage power equipment.

The AE signal caused by PD is received by the optical fiber sensor after passing through the center conductor of the basin-type insulator, the insulation structure with epoxy resin as the main component, and the metal shell. The birefringence index of PMF is affected by the pressured shock wave, resulting in a mapping relationship between the output light intensity and the AE signal intensity. By adjusting the distance between the discharge needle and the basin-type insulator, PD signals with different discharge intensities are obtained.

The collected signal is shown in Fig. 2. It can be clearly observed that the ultrasonic signal caused by PD is received by the optical fiber sensor after passing through the central conductor of the basin insulator, the insulation structure with epoxy resin as the main component, and the metal shell. The birefringence of PMF is affected by external force, resulting in obvious change of output light intensity. In this study, six different discharge modes are set up, the discharge distance is from 1 mm to 6 mm, the discharge voltage is 9 kV, and the sampling frequency is 1 MHz. The sampling duration of each discharge mode is 200 second, and the collected signals are split to 200 groups for training and verification of the SVM model during pattern recognition.

III. EIGENVALUES EXTRACTING

First, the time-domain and frequency-domain eigenvalues of the collected AE signal are optimized as input vectors to train the SVM model, and then the test set is used for verification. 16 time-domain eigenvalues are selected for inspection, namely, maximum, minimum, mean, absolute mean, peak, peak-to-peak, variance, standard deviation, effective value, square root amplitude, slope, kurtosis index, waveform factor, peak factor, pulse index and margin index; 3 frequency-domain eigenvalues are selected for inspection, namely, maximum power spectrum, median frequency and mean power frequency. Ultimately, 12 eigenvalues with accuracy more than 50% are selected by the classification accuracy of each eigenvalue, which is composed of 10 time-domain eigenvalues and 2 frequency-domain eigenvalues. The time-domain eigenvalues include maximum, minimum, mean, absolute mean, peak, effective value, square root amplitude, peak factor, pulse index and margin index. The frequency domain eigenvalues include the maximum power spectrum and the median frequency.

Second, the MFCC is extracted from the collected signals to train the SVM model. Before extracting MFCC, it is
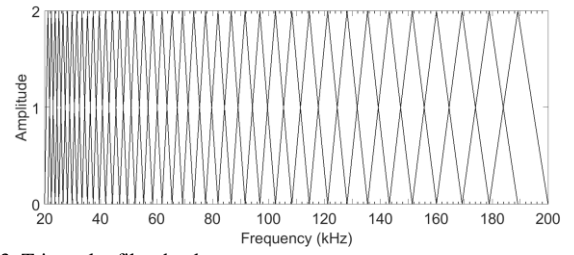
Fig. 3. Triangular filter banks.

necessary to establish a set of mel scale triangular filter banks. The relationship between Mel frequency and linear frequency is as follows: $Mel(f) = 1125 \times \ln(1 + f/700)$. The effective bandwidth of the optical fiber AE sensor developed by our research group in the previous work is 0-200 kHz, and the analysis frequency of the AE signal is greater than 20 kHz. Therefore, the filter frequency range here is 20-200 kHz, and the corresponding range on the Mel scale is 3810.16-6365.80. Here, 40 filters are selected in this frequency range, and 42 points are needed for this purpose, which means that we need to linearly interval 40 additional points between 3810.16 and 6365.80. The first filter bank will then start at the first point, peak at the second point, and return zero at the third point. The second filter bank will start from the second point, reach the maximum value at the third point, and then return zero at the fourth point, and so on , a set of mel scale overlapping triangular filter banks can be obtained, as shown in Fig. 3.

The next step is MFCC extraction. In MFCC extraction, N sampling points are first assembled into one observation unit, which is called frame, and the value of N covers 20 ms in this study. In order to avoid the large change of two adjacent frames, there is an overlapping area between the two adjacent frames, which contains M sampling points. Here, M is equal to N/2. Then multiply each frame by the Hamming window to increase the continuity of the left and right ends of the frame. A fast Fourier transform is performed on each frame of the framed windowed signal to obtain the frequency spectrum of each frame, and the signal spectrum is modulo-squared to obtain the signal power spectrum. The power spectrum is passed through the mel scale triangular filter banks shown in Fig. 3, and then the logarithmic energy of each filter bank is calculated. Finally, the above logarithmic energy is brought into the discrete cosine transform to obtain the MFCC parameters. After comparison and analysis, 13 dimensional MFCC parameters are selected in this study.

IV. PATTERN RECOGNITION AND ANALYSIS

For small-sample size or non-normal data, SVM has obvious training effect and high recognition accuracy, which is suitable for data analysis of small probability events such as PD fault of GIS equipment. Therefore, this study uses SVM model to recognize PD fault of GIS equipment. In this study, six kinds of PD discharge categories are designed, which belongs to multi classification problem. In the process of recognition, firstly, the feature value of the signal is extracted, then the SVM model is trained by training set based on the extracted feature values, finally the classification effect is verified by test sets. Here, 12 eigenvalues and MFCC are selected as input vectors of the SVM model. 200 sets of feature vectors of 6 kinds of simulated PD discharge categories are extracted from the collected signals. After the normalization of feature vectors, 180 sets of feature vectors are used as training samples, 20 sets of feature vectors are used as test samples, and the SVM model is trained and tested. The

Fig. 4. Classification and prediction results of test sets. (a) 12 eigenvalues, (b) MFCC, (c) MFCC+12 eigenvalues.

Discharge distance	12 eigenvalues	MFCC	MFCC+12 eigenvalues
1 mm	100%	85%	100%
2 mm	25%	45%	65%
3 mm	90%	100%	100%
4 mm	25%	75%	85%
5 mm	75%	100%	90%
6 mm	100%	100%	100%
Total	67.5%	84.17%	90%

frequency domain eigenvalues, the classification accuracy can reach 67.5%. Using MFCC feature vectors as SVM input vector for pattern recognition, the classification accuracy can reach 84.17%, which shows that MFCC eigenvalues have better recognition effect than 12 typical time-frequency domain eigenvalues. When MFCC and the 12 time-frequency domain eigenvalues are used as the input vectors of SVM, the recognition accuracy can reach 90%. It can be seen that after adding MFCC eigenvalue as the input feature vector of SVM, the overall recognition accuracy is improved from 67.5% to 90%, which is improved by one third, and the improvement effect of accuracy is very obvious. Among them, the recognition accuracy of 2 mm discharge type is improved by 40%, and that of 4 mm discharge type is increased by 60%.

Moreover, it should be emphasized that MFCCs are extracted from the original signal in this study, not the denoised signal. However, due to the complex structure of the basin insulator wound by the sensor, large attenuation and absorption of epoxy resin, electrical and mechanical noise and other influencing factors, the collected signals are relatively weak and the signal-to-noise ratio is very low, which affects the final recognition accuracy to a certain extent. In the next step, in order to further improve the accuracy of discharge pattern recognition, we will try to denoise the signals and then use them for pattern recognition. However, for such weak PD signals with low signal-to-noise ratio, the recognition accuracy of the proposed research method and solution can still reach 90%, which can better show the contribution of MFCC eigenvalues to the improvement of recognition accuracy. Therefore, MFCC is an effective characteristic value, which can be well used for online monitoring and analysis of PD faults in high-voltage equipment.

VI. CONCLUSION

In this study, the pin-to-plate discharge experiment platform is used to simulate the PD phenomenon in GIS equipment, and the signal is collected based on the optical fiber AE sensor. The typical time-domain and frequency-domain eigenvalues combined with Mel frequency cepstrum coefficient are used for SVM pattern recognition, and the recognition accuracy of the simulated PD fault can reach 90%. The research results show that the proposed research methods and solutions can greatly improve the accuracy of GIS internal closed environment defects prediction, which can provide technical support for the safe and reliable operation of high-voltage power supply system, and have better economic value and social significance.

experimental results verify the contribution of 12 time-frequency domain eigenvalues combined with MFCC as the input vector of SVM model.

The classification and prediction results of test sets are shown in Fig. 4. For each type of PD signal, 200 groups of eigenvectors are obtained, 180 groups are taken as training samples, 20 groups are taken as test samples, and there are 120 groups of test samples in 6 types of discharge signals in total. In the figure, abscissa represents 120 sets of test set eigenvectors, and ordinate represents 6 category labels, corresponding to 1 mm to 6 mm discharge distance. The blue circle represents the actual discharge type, and the red asterisk represents the discharge type predicted by SVM model. If the red asterisk coincides with the blue circle, it means the prediction is correct. On the contrary, if the red asterisk deviates from the blue circle, the prediction is wrong.

Table 1 shows the recognition accuracy results of each discharge type and the overall recognition accuracy results obtained by using different eigenvalues as input vectors of SVM. After training the SVM model with 12 typical time-

REFERENCES

[1] P. Bolin and H. Koch, "Introduction and applications of gas insulated substation (GIS)," IEEE Power Engineering Society General Meeting IEEE, 2005.

[2] J. Wang, W. Ding, Y. Liu, Z. Zheng, and C. Ge, "Application of X-Ray Inspection for Ultra High Voltage Gas-Insulated Switchgear," IEEE Transactions on Power Delivery, vol. 34 , pp. 1412-1422, 2019.

[3] K. Sasamori and H. Hama, "History of Gas-insulated Switchgear in Japan on Transition of Technological Development and Practical Application—The Third Report: Expanded Application of UHV and HVDC Technologies to GIS—," IEEJ Transactions on Fundamentals and Materials, 2019.

[4] Q. Khan, S.S. Refaat, H. Abu-Rub, and H.A. Toliyat, "Partial discharge detection and diagnosis in gas insulated switchgear: State of the art," IEEE Electrical Insulation Magazine, vol. 35, pp. 16-33, 2019.

[5] X. Li, Y. Zhang , H. Cao , X. Zhang, and R. Shen, "Highly Sensitive Acoustic Emission Sensor Based on Polarization-maintaining and Absorption-reducing Polarization-maintaining Fiber", Sensors and materials, vol.30, pp. 1145-1153, 2018.

[6] J. Guo, S. Zhang, L. Wei, " Ultrasonic Detection of Partial Discharge on Typical Defects of Basin Insulator in GIS," 2019 8th International Conference on Power Science and Engineering (ICPSE), 2019.

[7] H. Zhao, "A Novel Sensor Detection Method of Partial Discharge in GIS based on the UHF Electromagnetic Wave," IOP Conference Series Earth and Environmental ence, 2020.

[8] Q. Bo, G. X. Zhang, C. R. Li, C. J. Gao, and Z. Z. Chen, "Research status and prospect of gas-insulated metal enclosed transmission line," High Voltage Engineering, vol. 41, pp. 1466-1473, 2015.

[9] T. Shanker, H. N. Narasimhaiah, and G. Punekar, "Acoustic emission partial discharge detection technique applied to fault diagnosis: Case studies of generator transformers," Serbian Journal of Electrical Engineering, 2016.

[10] D. Ding, W. Gao, and W. Liu, "Insulation defects discrimination in GIS by fisher discriminant analysis of partial discharge," Gaodianya Jishu/high Voltage Engineering, vol. 39, pp. 805-812, 2013.

[11] W. Yang, et al, "GIS Partial Discharge Patterns Recognition with Spherical Convolutional Neural Network," 6th International Conference on Electrical Engineering Control and Robotics (EECR 2020), 2020.

[12] X. Li, X. Wang, A. Yang, and M. Rong, "Partial Discharge Source Localization in GIS Based on Image Edge Detection and Support Vector Machine," IEEE Transactions on Power Delivery, 2019.

[13] Y. Ling, D. Bai, M. Wang 2, X. Gong, and C. Gu, "SVM-based Partial Discharge Pattern Classification for GIS," 2017 2nd International Conference on Communication, Image and Signal Processing (CCISP 2017), 2017.

[14] S. B. Davis, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," IEEE Trans, Acoust, Speech Signal Process, vol. 28, pp. 65-74, 1980.

[15] W. L. Shi, and X. H. Fan, "Research on armored vehicle classification based on MFCC and SVM," IEEE International Conference on Computer & Communications IEEE, 2018.

[16] J. Gao, Y. Zhu, Y. Jia, Y. Zheng, and S. Liu, "Pattern Recognition of Unknown Types of Partial Discharge Based on Improved SVDD Algorithm and Mahalanobis Distance," Diangong Jishu Xuebao/Transactions of China Electrotechnical Society, vol.33, pp. 3510-3517, 2018.

[17] Q. Q. Zhang, J. Lin, H. Song, and G. H. Sheng, "Fault Identification Based on PD Ultrasonic Signal Using RNN, DNN and CNN," Condition Monitoring and Diagnosis, 2018.

[18] Q. Che, H. Q. Wen, X. Y. Li, Z. Q. Peng, and K. P. Chen, "Partial Discharge Recognition Based on Optical Fiber Distributed Acoustic Sensing and a Convolutional Neural Network," IEEE The Institution of Engineering and Technology, vol. 7, pp. 101758-64, 2019.

[19] Z. X. Wei, Y. Ju, and M. Song, "A Method of Underwater Acoustic Signal Classification Based on Deep Neural Network," 2018 5th International Conference on Information Science and Control Engineering (ICISCE) IEEE, 2018.

[20] B. Yan, G. Q. Qian, F. H. Wang, and S. Chen, "Noise recognition of power transformers based on improved MFCC and VQ," 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D) IEEE, 2016.

[21] H. Qin, W. Y. Zhou, M. Z. Zhang, and P. X. Liu, "Research on fault diagnosis method of distribution transformer based on MFCC and HMM," Proceedings of the 2016 4th International Conference on Sensors, Mechatronics and Automation (ICSMA 2016), 2016.

Muscle Temperature Sensing and Control with a Wearable Device for Hand Rehabilitation of People After Stroke

Han Li
School of Instrument Science and Engineering
Southeast University
Nanjing, China
hanli@seu.edu.cn

Jun Zhang*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
j.zhang@seu.edu.cn

Chaojun Jiang
School of Instrument Science and Engineering
Southeast University
Nanjing , China
984212445@qq.com

Qi Liu
School of Instrument Science and Engineering
Southeast University
Nanjing, China
220193279@seu.edu.cn

Maozeng Zhang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
354926364@qq.com

Jingsong Zhou
School of Instrument Science and Engineering
Southeast University
Nanjing, China
alonappy@163.com

Abstract—**Muscle spasm affects the hand rehabilitation of the person after stroke. This paper presents the muscle temperature sensing and control with a wearable device. The device mainly consists of three layers, i.e. a Graphite heat dissipation film, a Peltier array, tailor-made radiation fins, from bottom to top. Multiple temperature sensors PT1000 installed between the film and the Peltier pieces are employed for distributed detecting the surface temperature of the muscle which drives the motion of fingers. The Peltier array is used to control the temperature of the muscle with a PID controller for regulating the voltage supplied for the Peltier array. The direction of the current through the Peltier array can be adjusted for cold and heat stimulation of the muscle. By precisely controlling the temperature of the muscle, this device could alleviate the muscle spasm and reduce the edema of hand for better rehabilitation treatment after stroke. The device can also be used for investigating other symptoms alleviation needing cold stimulation or thermotherapy.**

Keywords—Peltier, detecting, PID, stimulation, rehabilitation

I. Introduction

Stroke is one of the main causes of disability in the aging society worldwide [1-2]. The number of stroke people grows more and more in the recent years. Stroke leads to hemiplegia and loss of some motor functions of the patients. This not only results in great difficulties to their personal lives, but also increases the burden on families and society [3]. Hand motor dysfunction is one of the common symptoms, whose recovery is slow and difficult because hand movement is controlled by a large number of nervous systems.

Timely rehabilitation treatment is very important for the recovery of motor function of the stroke patients. There are many methods for the rehabilitation of stroke patients, including the exercise therapy [4-6], neuromuscular electrical stimulation [7-8], transcranial magnetic stimulation [9], cold therapy, mental practice [10], thermotherapy, Warm-needle moxibustion [11], etc.

Muscle spasm and hand swelling are complications of the stroke patients causing functional impairment and pain to the stroke survivors [12-13]. Muscle spasticity also limits the recovery of stroke survivors [14]. In the rehabilitation methods mentioned above, thermotherapy and cold therapy
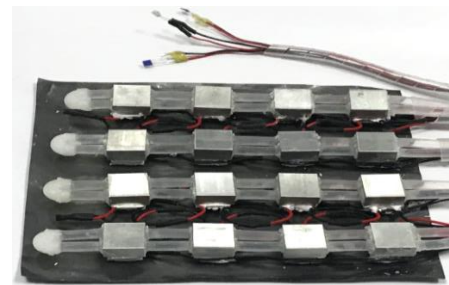
Fig. 1. Prototype of the wearable device with muscle temperature sensing and control capabilities to alleviate muscle spasm and swelling in hand rehabilitation.

could be used to alleviate the muscle spasticity and swelling, and relief the pain suffered by the patient. D. Uivarosan [1] employed thermotherapy to reduce spasticity of the stroke patients. C. R. Denegar [15] conducted cold therapy experiments with ice and found that cold is effective in treating the pain and muscle spasm. D. M. Daniel et al. [16] used cooling pads to lower the skin temperature to study the effects of cold therapy on the pain, swelling, and range of motion of the patient after surgery. R. W. Culp et al. [17] used ice bag and a cooling blanket for lowing skin temperature of people after surgery of the hand and wrist and found that the cooling blanket is more efficient and accurate than ice bags. E. Morsi [18] found continuous-flow cold therapy is beneficial for the patients after knee arthroplasty. E. A. Mailler-Savage [19] utilized cold therapy system to decrease pain and inflammation of the patient after knee operation. B. Altintas et al. [20] studied the influence of immediate cold therapy on burns to relief pain and limit tissue damage.

However, the ice bag usually used in the cold therapy has some disadvantages. The ice should be prepared before the treatment and filled into the bag at an interval during treatment because the ice melts in the bag. This is inconvenience in use. The melting of the ice also makes the temperature control be inaccurate [17]. Too low temperature may cause damage of the tissue while too high temperature may reduce the effect of the treatment. In the thermotherapy, the temperature of the warm water also changes with time. Some medical instruments are used for thermotherapy. However, they are usually used at hospital which is inconvenient for the rehabilitation of stroke patients.

---

* Jun Zhang is the corresponding author (e-mail: j.zhang@seu.edu.cn).

As shown in Fig. 1, in this paper, we present a design method for a wearable device, which can be used to accurately detect the surface temperature of muscles, and can precisely control the temperature. By cold stimulation, the device can alleviate muscle spasms. And the device can also reduce swelling and pain in patients by thermal stimulation. The device is beneficial to stroke patients in recovery treatment. The rest of this paper is organized as follows. Section II introduces the design of the device. The stimulation modes of the device and its temperature control method are presented in Section III. The prototype and experimental results are given in Section IV. The last section summarizes this work and gives the future direction of this study.

II. WEARABLE DEVICE DESIGN

The proposed wearable device is composed of a wearable subsystem, a water cooling subsystem and a temperature control subsystem. The wearable subsystem realizes the functions of temperature detection and muscle stimulation. The water cooling subsystem completes the heat dissipation function of the wearable subsystem. The temperature control system provides the power, drives the other subsystems, and control the stimulation temperature.

A. Wearable Subsystem

The structural diagram of the wearable subsystem is depicted in Fig. 2. It is mainly divided into three layers, including a thermal conductive graphite film, a Peltier array, and a heat sink array, from bottom to top. The graphite film is closely attached to the muscles so as to achieve efficient heat conduction. The Peltier pieces have a small volume and are arranged in an array. Hence, the uniform temperature control can be achieved as much as possible. A metal aluminum block is installed on each Peltier piece with water flowing inside to dissipate heat. In addition, PT1000 platinum resistors are installed between the graphite film and the Peltier pieces, and are used to detect the muscle surface temperature of the patients. What's more, the Peltier pieces are cascaded in series, which is convenient for designing a drive control system and realizing drive control in different regions. The heat sinks are connected with 3D printed joints and rubber water pipes. The heat sinks connect with the water cooling subsystem to achieve heat dissipation of the Peltier pieces.

(a)   (b)
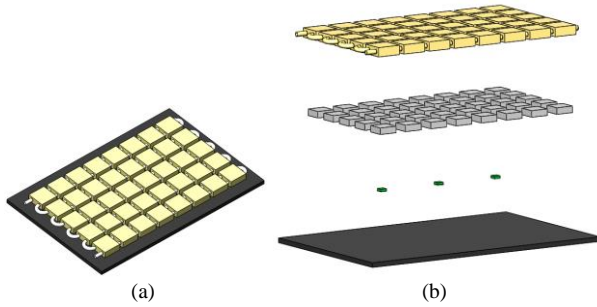
Fig. 2. Structural diagram of the wearable subsystem with three layers. (a). Cross-sectional view. (b). Exploded view.

TABLE I   PARAMETERS OF THE PELTIER USED IN THERMODYNAMIC SIMULATIONS.

Size	Maximum pump heat	Maximum current	Maximum voltage
8.3mm×8.3mm ×2.8mm	2.5W	1.2A	3.66V

B. Cooling and Control Subsystems

There are mainly two methods usually used for heat dissipation. One is airflow cooling and the other one is water cooling. In order to select a proper cooling method for achieving good cooling performance, the thermodynamic simulation analysis of the cooling effect of the Peltier array, without active cooling, with the airflow cooling, and with the water cooling is carried out in this paper. The parameters of the Peltier pieces used are shown in Table I.

The simulation results are shown in Fig. 3, which shows the thermal load distribution in the three conditions. The change of the temperature at the cold side of the Peltier piece with time in the three conditions are shown in Fig. 4. It can be seen that, compared with no active heat dissipation, airflow cooling and water cooling have obvious effects on stabilizing the temperature. In the case of airflow cooling, the temperature of the cold side of the Peltier will gradually increase with time, while the temperature of the cold side of the Peltier in the case of water cooling is well maintained at a low temperature. The thermodynamic simulation analysis shows that the water cooling is better than no active cooling and air cooling. Therefore, we use a water cooling system to achieve active heat dissipation of the heat sink.

(a)   (b)

(c)
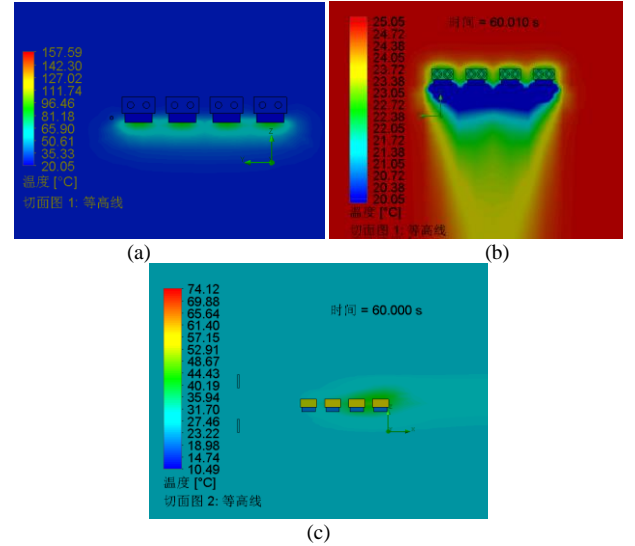
Fig. 3. The thermodynamic simulation results of the cooling system with different cooling methods. (a) Without active cooling. (b) Airflow cooling. (c) Water cooling.
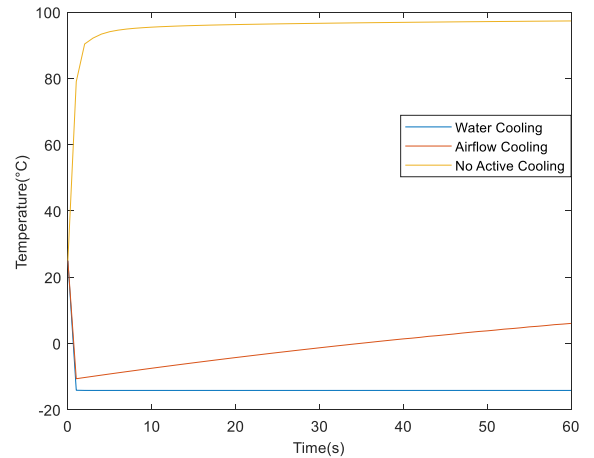
Fig. 4. Comparison of the temperature changes of cold side of the Peltier piece in the three conditions.
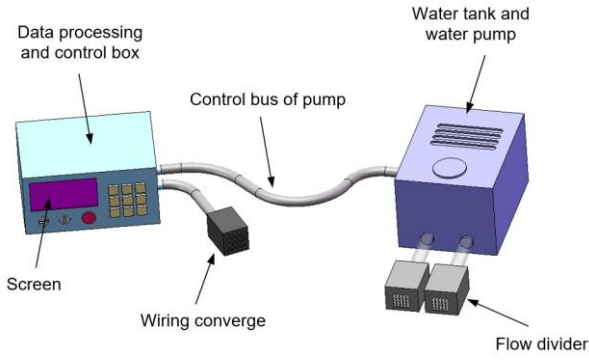
Fig. 5. The water cooling subsystem and temperature control subsystem.

Fig. 6. Flow of temperature detection and control system.

As shown in Fig. 5, the water cooling subsystem includes a water tank, a water pump, flow dividers, and water pipes. In order to control the Peltier arrays in different areas to achieve precise temperature stimulation, we designed a temperature control subsystem. The control subsystem is mainly composed of a power supply drive module, a data acquisition module, a MUC board, a human-computer interaction interface, and an electrical interface. The power supply driving module adopts a multi-channel power supply module, which sets the output voltage through serial port communication, thereby realizing precise temperature regulation. The control subsystem also controls the start-stop of the water cooling subsystem.

The data acquisition module will convert the resistance of the wearable subsystem temperature sensor PT1000 through a dedicated temperature measurement conversion chip MAX31865 to obtain the temperature information. Then the MCU runs data processing and temperature control algorithms to achieve precise temperature control. The human-computer interaction interface allows users to set the stimulation mode, stimulation temperature, and other parameters. The electrical interface is used for power supply and communication with the wearable subsystem and the water cooling subsystem.

III. TEMPERATURE CONTROL FOR COLD AND HEAT STIMULATIONS

The muscle spasm and hand edema treatments need control of the muscle temperature in cold stimulation and thermotherapy. Accurate and stable temperature control is the key to achieving a good treatment. This section analyzes the temperature stimulation modes of the wearable device and introduces the methods of temperature acquisition and control.

A. Stimulation Patterns

Because cold therapy and thermotherapy require different temperature for muscle stimulations. For example, the temperature required for cold stimulation is about 0°C to 15°C while the temperature required for thermal stimulation is about 45°C to 60°C. The wearable device designed in this paper adopts a modular design concept. We can splice multiple wearable units into a large integrated device according to the size of the limb muscles to be stimulated. The proposed wearable device has multiple stimulation modes.

(1) Cooling or heating: The wearable device adopts the Peltier as the core component. Cooling and heating can be realized only by changing the direction of the current through the Peltier array.

(2) Stimulation on different regions: The wearable device currently designed has two channels. Each channel works independently without interfering with each other. According to the different rehabilitation strategies of cold therapy and thermotherapy, targeted stimulation treatment can be carried out for different regions of the limbs, which increases the accuracy of treatment in rehabilitation and may improve the recovery efficiency.
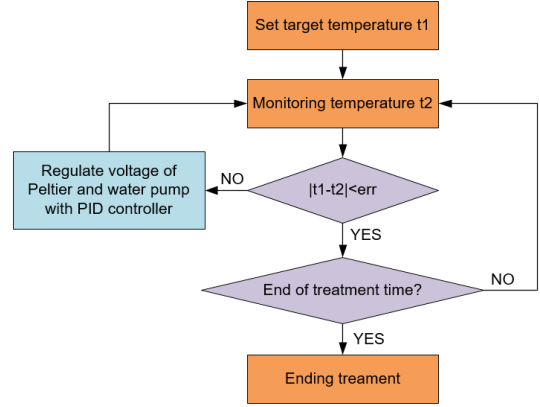
(3) Stimulation at different times: The multiple channels of the wearable device can not only realize individual stimulation on different regions, but also achieve different choices of stimulation duration. The time-sharing stimulation greatly increases the options of cold therapy and thermotherapy, which is of great significance for the rehabilitation of patients.

Combining with the above three kinds of stimulation patterns, the wearable device can realize different stimulations on different regions and at different time with intelligent control. We call it a 4D stimulation method which has flexible stimulation mode in the therapy and can meet various application needs, such as periodic stimulation therapy for cold stimulation and thermal stimulation.

B. Temperature Detection and Control

In order to achieve accurate 4D cold and thermal stimulations, the multi-point temperature detection method is used in this paper. Real-time temperature monitoring is performed by installing a temperature sensor on the lower surface of each Peltier piece. There are several temperature measurement probes, which can detect the temperature of the part on the limb where no cold or thermal stimulation is performed, or the temperature of the healthy limb, as a reference, for comparison and analysis of the stimulations.

The control of the device and the muscle temperature includes two parts. Firstly, the collected temperature and the set stimulation temperature are compared. A PID controller is used to adjust the power supply voltage of the Peltier array to achieve real-time temperature control. Secondly, it is necessary to control the rotation speed of the water pump in the cooling process, so as to realize the water cooling and adjust the cooling effect. The system control flow is shown in Fig. 6. The system adopts different control strategies according to the stimulation mode selection of users.

IV. EXPERIMENTS

In order to test the effect of the wearable device proposed in this paper and verify the feasibility of the proposed cold and thermal stimulation for hand rehabilitation, we designed a prototype system and conducted several performance experimental tests in this section.
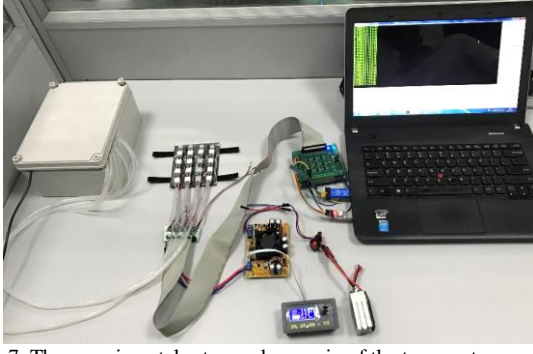
Fig. 7. The experimental setup and scenario of the temperature control experiments.

A. Prototype

Fig. 1 is a photo of the prototype of the wearable device. The wearable device has a 4×4 Peltier array, which is divided into 2 channels, each with a 2×4 Peltier array. The parameters of the Peltier are shown in Table I. And there is a PT1000 temperature sensor between the thermal conductive graphite film and each Peltier. Besides, the heat sink is made of aluminum alloy and the size is 10mm×10mm×5mm. And the gap between adjacent heat sinks is 10mm, so the size of the entire wearable subsystem is about 70mm×70mm×9mm. The system implemented for the tests is shown in Fig. 7.

The power lines of Peltier and the signal lines of the temperature sensors are connected to the control system through an adapter and a cable. The water pipe of the heat sink and the water tank are connected by a shunt. What's more, we designed the data acquisition and control circuit board based on STM32F103ZET6 MCU, and realized the software of data acquisition and temperature control algorithm. In addition, 3D printing was used to complete the production of the water tank and connectors, and the connection with the wearable device also was completed.

B. Performance of the Device

a) Comparison of cooling and heating effects with and without water cooling

In this experiment, the room temperature is controlled at about 25°C, and the target temperature is set at 10°C. Then conduct the experiments with water cooling and without active cooling, and collect the data from 16 temperature sensors. Finally, calculate the average value and draw the temperature-time curve as shown in Fig. 8.
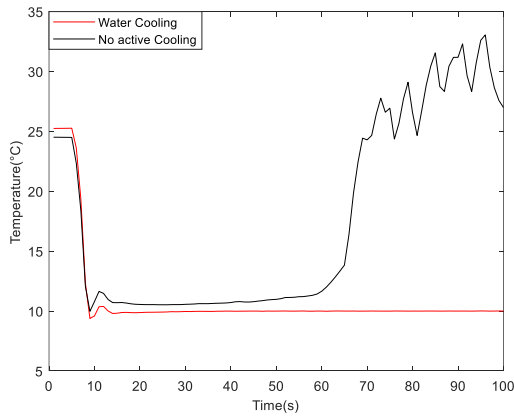
Fig. 8. Average temperatures detected by the sensors with water-cooling and without active cooling.

(a)

(b)

(c)

(d)
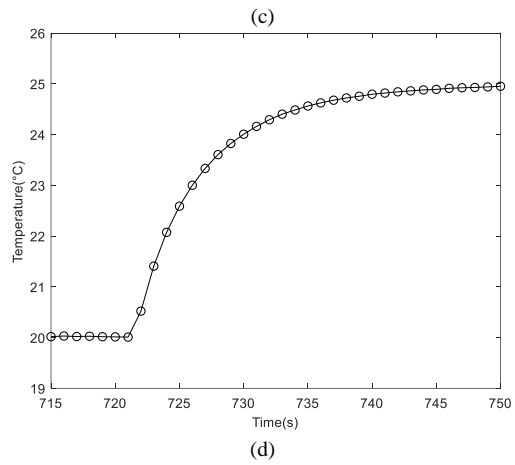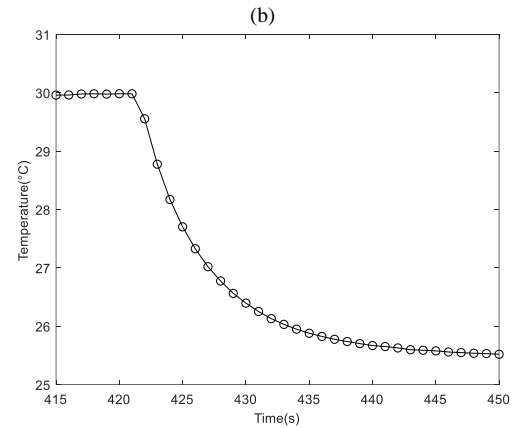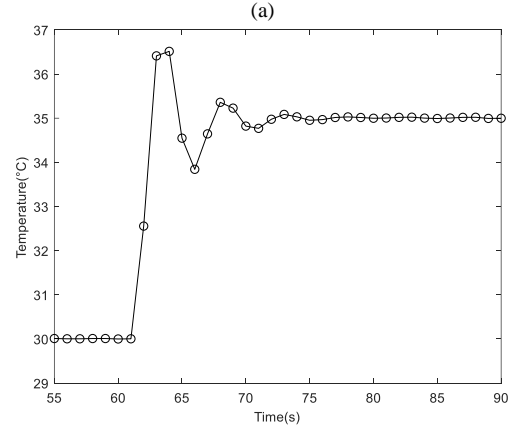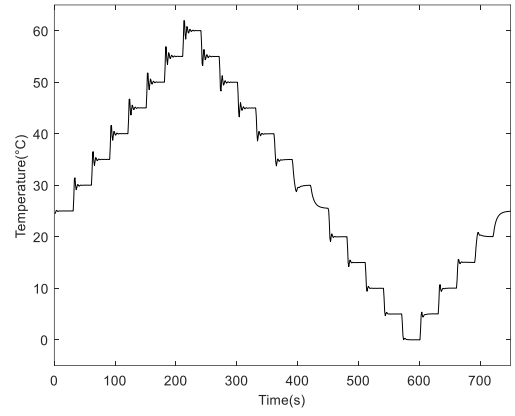
Fig. 9. Temperature control performance. The temperature increases from 25°C to 60°C firstly, and then decreases to 0°C, and increases to 25°C finally. The changing step is set as 5°C. (a) Overall process. (b) Process of 30°C→35°C. (c) Process of 30°C→25°C. (d) Process of 20°C→25°C.

As can be seen from Fig. 8, without active cooling, Peltier cannot maintain a constant temperature for a long time, and its temperature will gradually increase with time. On the contrary, Peltier with water cooling can achieve good temperature control. This result also corresponds to the simulation result. Water has a large specific heat capacity, which is enough for the heat dissipation of the system. And compared with other liquids, the advantage of water is that it can be acquired almost everywhere, so it is easy to replace and beneficial to the portability of the system. Therefore, water cooling is selected.

b) Accuracy, repeatability and response speed of temperature control

Place the Peltier array in an environment with a room temperature of 25°C. After its temperature is stable, control the temperature change process of the Peltier array to be 25→30→35→40→45→50→55→60→55→50→45→40→35→30→25→20→15→10→5→0→10→15→20→25 (Unit: °C), and measure the data of 16 temperature sensors.

Fig. 9. shows the average temperatures of 16 sensors, where (a) is the temperature-time curve of the entire experiment and (b) is the image of the process from 30°C to 35°C. It can be seen that Peltier's response time is about 7s, and its repeatability and accuracy are very high.

However, the system has some deficiencies. When the temperature is from 30°C to 25°C (as Fig. 9(c)) or from 20°C to 25°C (as Fig. 9(d)), the heat dissipation efficiency is very low because the temperature is very close to the water temperature, which is 25°C. Therefore, the response speed becomes slow.

This deficiency seems to have little effect on system performance, because the temperature required for cold stimulation is about 0°C to 15°C while the temperature required for thermal stimulation is about 45°C to 60°C, in which the system has an excellent performance.

c) Long-term stability of temperature control

In order to test the stability of the system for a long time, the device is placed in an environment with a temperature of 25°C, and work for 30 minutes with the target temperature of 10°C. The average temperature is shown in Fig. 10.

It can be seen from Fig. 10 that the equipment can quickly reach the target temperature and can remain stable for a long time, which shows that the device has excellent long-term stability.
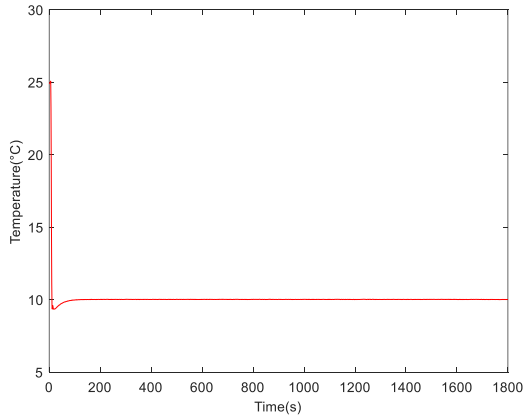
Fig. 10. Average temperatures detected by the sensors with water-cooling for long-term stability test.
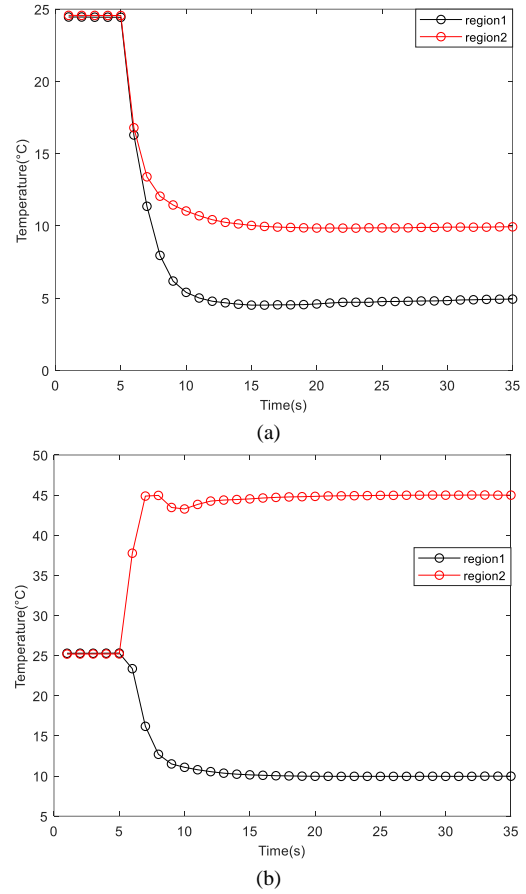
Fig. 11. The two regions reach different target temperatures. (a) The target temperature are 5°C and 10°C respectively. (b) The target temperature are 10°C and 45°C respectively.

d) Temperature control of different regions

The 4×4 Peltier array is simply and evenly divided into 2 regions, and each region has a 2×4 Peltier array. Set different target temperatures for the two regions at the same time, the average temperature of each region is shown in Figure 11.

According to the experimental results, this device can achieve independent stimulation on different areas, so that targeted stimulation treatment can be carried out for different regions of the limbs.

V. CONCLUSION

This paper presents the muscle temperature sensing and control with a wearable device. The device mainly consists of a Graphite heat dissipation film, PT1000 platinum resistors, a Peltier array and tailor-made radiation fins. The Peltier array is used to control the temperature of the muscle with a PID controller, which regulates the voltage supplied for the Peltier array. After simulation and experimental verification, the effect of temperature control with water cooling is far stronger than that without active heat dissipation, and the Peltier array in this paper can realize multiple stimulation modes with high accuracy and fast response speed. Therefore, this device could be used to precisely control the temperature of the muscle for better rehabilitation treatment after stroke.

However, this system also has some shortcomings. So the future research directions mainly include the following points:

(1) The current arrangement and regional division of Peltier is very simple. In the future, the number, arrangement and regional division of Peltier will be planned according to the distribution of limb muscles.

(2) When the temperature of Peltier is close to the temperature of water, the speed of temperature adjustment slows down due to the decrease in the heat dissipation rate. In the future, some algorithms may be used to change the voltage at this time, so as to change the heat generation or cooling rate of the Peltier, and achieve the purpose of accelerating the temperature response rate.

(3) In the future, patients' data will be collected to verify the effectiveness of the system and to improve the method.

REFERENCES

[1] D. Uivarosan, D. Mirela Tit, C. Iovan, D. C. Nistor-Cseppento, L. Endres, "Effects of combining modern recovery techniques with neurotrophic medication and standard treatment in stroke patients," *Science of the Total Environment*, vol. 679, pp. 80-87, Aug. 2019.

[2] H. Wang, C. J. L. Murray, A. Lopez, "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015," *Lancet*, vol. 388, no. 10053, pp. 1459-1544, 2016.

[3] G.J. Hankey. "The global and regional burden of stroke," *Lancet Glob. Health*, vol. 1, no. 5, pp. e239-e240, Nov. 2013.

[4] K. Triandafilou, D. Kamper, "Carryover effects of cyclical stretching of the digits on hand function in stroke survivors," *Arch. Phys. Med. Rehab.*, vol. 95, pp. 1571-1576, 2014.

[5] E. Ghasemi, K. Khademi-Kalantari, M. Khalkhali-Zavieh, A. Rezasoltani, M. Ghasemi, "The effect of functional stretching exercises on functional outcomes in spastic stroke patients A randomized controlled clinical trial," *Journal of Bodywork & Movement Therapies*. vol. 22, no 4, pp. 1004-1012, 2018.

[6] K. Dworzynski, G. Ritchie, E. D Playford, "Stroke rehabilitation : long-term rehabilitation after stroke," *Clinical Medicine*, vol.15, no 5, pp. 461-464, 2015.

[7] AI Bello, BE Rockson, MO Olaogun, "The effects of electromyographic triggered neuromuscular electrical muscle stimulation on the functional hand recovery among stroke survivors," *Afr J Med Med Sci*, 2012.

[8] Y. Hara, S. Obayashi, K. Tsujiuchi, Y. Muraoka, "The effects of electromyography-controlled functional electrical stimulation on upper extremity function and cortical perfusion in stroke patients," *Clinical Neurophysiology*, vol. 124, pp. 2008–2015, 2013.

[9] V.P. Motamed, F. Bahrpeyma, M. Firoozabadi, "Low Frequency repetitive Transcranial Magnetic Stimulation to Improve Motor Function and Grip Force of Upper Limbs of Patients With Hemiplegia," *International Journal of the Iranian Red Crescent Society*, 2014, 16(8) : e13579-e13579.

[10] H. T. Peters and S. J. Page, "Integrating Mental Practice with Task-specific Training and Behavioral Supports in Poststroke Rehabilitation: Evidence, Components, and Augmentative Opportunities," *Physical Medicine and Rehabilitation Clinics of North America*, vol. 26, no. 4, pp. 715-727, Nov. 2015.

[11] L. Yang, J. Tan, H. Ma, H. Zhao, J. Lai, "Warm-needle moxibustion for spasticity after stroke: a systematic review of randomized controlled trials," *International Journal of Nursing Studies*, vol. 82, 2018, pp. 129-138.

[12] A. Opheim, A. Danielsson, M. Alt Murphy, H.C. Persson, K.S. Sunnerhagen, "Upper-limb spasticity during the first year after stroke: stroke arm longitudinal study at the University of Gothenburg," *Am. J. Phys. Med. Rehabil. 93*, pp. 884–896. 2014.

[13] Wissel, J., Ganapathy, V., Ward, A.B., Borg, J., Ertzgaard, P., Herrmann, C, "Onabotulinumtoxin A improves pain in patients with post-stroke spasticity: findings from a randomized, double-blind, placebo-controlled trial," *J. Pain Symptom Manag.* 52, 17–26.

[14] X. Hu, N. L. Suresh, M. K. Chardon, W. Z. Rymer, "Contributions of motoneuron hyperexcitability to clinical spasticity in hemispheric stroke survivors," *Clin. Neurophysiol.* vol. 126, no. 8, pp. 1599-1606, Aug. 2015.

[15] C. R. Denegar and D. H. Perrin, "Effect of Transcutaneous Electrical Nerve Stimulation, Cold, and a Combination Treatment on Pain, Decreased Range of Motion, and Strength Loss Associated with Delayed Onset Muscle Soreness," *J Athl Train*, vol. 27, no. 3, pp. 200-206, 1992.

[16] D. M. Daniel, M. L. Stone, D. L. Arendt, "The effect of cold therapy on pain, swelling, and range of motion after anterior cruciate ligament reconstructive surgery," *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, vol. 10, no. 5, pp. 530-533, Oct. 1994.

[17] R. W. Culp and J. S. Taras, "The Effect of Ice Application versus Controlled Cold Therapy on Skin Temperature When Used with Postoperative Bulky Hand and Wrist Dressings: A Preliminary Study," *Journal of Hand Therapy*, vol. 8, no. 4, pp. 249-251, 1995.

[18] E. Morsi, "Continuous-Flow Cold Therapy After Total Knee Arthroplasty," *The Journal of Arthroplasty*, vol. 17, no. 6, 718-722, 2002.

[19] E. A. Mailler-Savage and D. F. Mutasim, "Cold injury of the knee and lower aspect of the leg after knee surgery and use of a cold therapy system," *Journal of the American Academy of Dermatology*, vol. 58, no. 5, Supplement 1, pp. s106-s108, May 2008.

[20] B. Altintas, A.A. Altintas, R. Kraemer, H. Sorg, P.M. Vogt, "Acute effects of local cold therapy in superficial burns on pain, in vivo microciculation, edema formation and histomorphology," *Burns*, vol. 40, no. 5, pp. 915-921, Aug. 2014.

Action-driven Reinforcement Learning for Improving Localization of Brace Sleeve in Railway Catenary

Junping Zhong
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
zhongjunping@my.swjtu.edu.cn

Zhigang Liu*
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
liuzg_cd@126.com

Hongrui Wang
Section of Railway Engineering
Delft University of Technology
Delft, Netherlands
H.Wang-8@tudelft.nl

Wenqiang Liu
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
Liuwq_2009@126.com

Cheng Yang
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
yangc@my.swjtu.edu.cn

Alfredo Núñez
Section of Railway Engineering
Delft University of Technology
Delft, Netherlands
a.a.nunezvicencio@tudelft.nl

Abstract—**Brace Sleeve (BS) plays an essential role in connecting and fixing cantilevers of railway catenary systems. It needs to be monitored to ensure the safety of railway operations. In the literature, image processing techniques that can localize BSs from inspection images are proposed. However, the boxes produced by existing methods can contain incomplete and/or irrelevant information of the localized BS. This reduces the accuracy of BS condition diagnosis in further analyses. To address this issue, this paper proposes the use of an action-driven reinforcement learning method that adopts the coarse-localized box provided by existing methods, and finds the movements needed for the box to approach to the true BS position automatically and accurately. In contrast to the existing methods that predict one position of the box containing a BS, the proposed action-driven method sees the localization problem as a dynamic position searching process. The localization of BS is achieved by following a sequence of actions, which in this paper are position-moving (up, down, left or right), scale-changing (scale up or scale down) and shape-changing (fatter or taller). The policy of selecting dynamic actions is obtained by reinforcement learning. In the experiment, the proposed method is tested with real-life images taken from a high-speed line in China. The results show that our method can effectively improve the localization accuracy for 81.8% of the analyzed images. We also analyze cases where the method did not improve the localization and suggest further research lines.**

Keywords—railway catenary, localization, brace sleeve, reinforcement learning, action-driven learning.

I. INTRODUCTION

Catenary is an important component of the traction power supply system in high-speed railways. A key component in catenary is the brace sleeve (BS). BS plays an important role in connecting and fixing catenary cantilevers. Due to the physical/mechanical impact triggered by the high-speed vehicles and other location and environmental factors along the railway line, the BSs can develop defective states. Defective BSs increase the risk of disrupting the railway operation and compromising safety. To automatically monitor the catenary components, image processing methods have been developed to replace manual checking. Once a defective component is detected, the information updates the maintenance activities planning so that the component can be replaced. The first step of monitoring is the localization of BSs. Localization is an important issue because if it is not accurate, fault detection methodologies will not count with the correct information to perform diagnosis [1].

*Zhigang Liu is the corresponding author. (e-mail:liuzgcd@126.com).

(a) Incomplete BS (b) Unnecessary information

(c) Complete BS (d) BS without unnecessary information
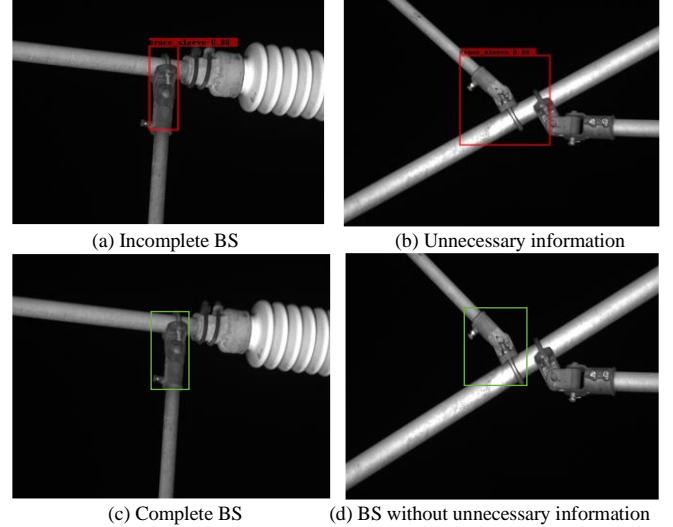
Fig. 1 Examples of detected BS using Faster R-CNN with a box containing (a) incomplete information, (b) unnecessary information. In (c) and (d) it is possible to see the optimal boxes for the previous cases.

In the past decade, two classes of localization algorithms have been widely used for railway component localization. The first class of algorithms are based on handcrafted features. Han *et al*. [2] used cascade support vector machines to classify a series of sliding window images, which are represented by a HOG to localize the catenary clevis. Zhong *et al*. [3] applied template matching on a standard catenary sleeve image and an original image to search the object position based on SIFT. Fan *et al*. [4] proposed a line LBP encoding method to represent a target object, which is used to localize fasteners on the rail track. The second class of algorithms are based on deep learning, which adopt supervised learning to train deep regression models that predict object position directly. In [5], a region-based convolutional neural network called Faster R-CNN is proposed to extract deep CNN features and localize general objects. Cai *et al*. [6] cascaded several regression modules in Faster R-CNN to further refine localization. Liu *et al*. [7, 8] and Kang *et al*. [9] applied improved Faster R-CNN to localize class-specific component, such as isoelectric line, brace sleeve screws and insulator. In works [10], deep learning architectures were developed to localize all catenary support components. Redmon *et al*. [11] introduced a strong deep CNN architecture called YOLO (You only look once) which allows to obtain a good trade-off between speed of detection and accuracy. Chen *et al*. [12] proposed an

improved YOLO for catenary components localization. Overall, handcrafted feature-based methods are simpler, but the performance of deep learning-based methods is by far superior for detecting catenary components. However, even state-of-the-art deep learning methods may provide incorrect localizations, either the localized BS is incomplete and/or with unnecessary information, as shown in Fig. 1. In this work, we propose a reinforcement learning (RL) method to address localization problems like the ones shown in Fig. 1.

The RL refers to a broad group of learning techniques. RL emulates the way living beings learn by trying actions and learning from successes and failures. As shown in Fig. 2, in RL, an agent is trained to make good decisions in a given environment by receiving rewards when the decisions are considered positive. The agent observes the state of a given environment, and takes actions that transform the environment to a new state according to its state-action policy, which is learned during training. A *Markov Decision Process* (MDP) is a formal mathematical representation of how the agent interacts with the environment to learn its policy. Recent works [13, 14, 15] in the RL field have proposed to combine deep neural networks with RL algorithms such as value function or policy function. By resorting of deep learning features, many difficult problems such as playing Atari games [17] or Go [13] can be successfully solved in a semi-supervised setting. For computer vision problems, various methods have been proposed in the literature. Caicedo *et al*. [18] proposed an active class-specific localization approach. Yun *et al*. [19] proposed an action decision method for object tracking by RL. In [20-21], RL was adopted to learn a policy of selecting a region from five fixed sub-regions, and realize object localization by only a few steps. So far, we are not aware of available literature applying RL to solve catenary component localization problems.
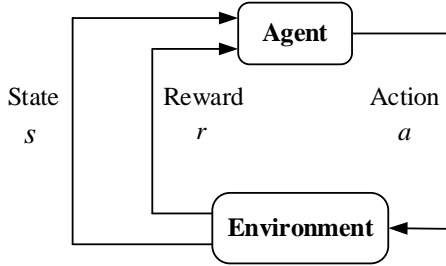
Fig. 2 A schematic flow of reinforcement learning system.

As shown in Fig.1(a) and Fig. 1(b), the boxes localized by existing region-based convolutional neural networks (R-CNN) methods do not enclose BS components tightly. In this paper, motivated by the reward-action in reinforcement learning and [19], we consider the localization improving problem as a control problem where a sequence of steps to refine the geometry of the localization box is to be obtained. Then, the localization refinement becomes a Markov Decision Process that can be trained with RL. We define the actions as position-moving, scale-changing and shape-changing. The reward is feedback about how well the current localization performed. Therefore, the action-decision policy can be learned according to the rewards being obtained. The agent is a deep CNN called ADNET (Action decision network) [19], which is presented in Section II. The application of reinforcement learning for BS localization is described in Section III. Experimental results and conclusions are given in Section IV and Section V, respectively.

The contributions of this paper are summarized as follows:

1. We investigate one possible method that employs RL to train an algorithm to generate a better bounding box for BS localization through a sequence of actions.

2. Different than the existing localization strategies for railway catenary systems, that localize objects following a single structured prediction model, the proposed method is a dynamic strategy that requires emphasis in the learning procedure and learning based on the time evolution of the performance (called history).

3. The preliminary results indicate that the proposed method is effective. The localization accuracy is improved while the time cost is low, which is beneficial for BS monitoring in railway.

II. LOCALIZATION IMPROVED BY REINFORCEMENT LEARNING

A. Method Overview

The overview of the action-driven method is shown in Fig. 3. The initial box image is an input for a deep CNN called ADNET (Action decision network). ADNET will select one of the actions, which are defined as transformations of the box by moving, scale changing, and shape-changing. Then, the initial geometry box is changed after taking the selected action and produces a new box image, which is sent to the ADNET to decide for the next action. Finally, the BS component is accurately localized by taking a sequence of actions. In this dynamic process, the applied control strategy of action selection is learned by reinforcement learning, which considers the performance feedbacks of all actions at each step. The learned action policy is aimed to automatically adjust the initial box to enclose the BS component automatically.

B. ADNET Structure

As shown in Fig. 3, the agent ADNET consists of three convolutional layers and four fully-connected layers. The parameters of conv1~conv3 and fc4 are similar to the widely used VGG [22] setting. The fc5 is concatenated by a base 1*1*512 and an action history vector. The fc6 is set as an action layer, whose output is a m-dimensions vector that represents probabilities of m actions. The fc7 is a 2-dimensions vector for classification (object/background). The input size is set to be 112*112*3. When an image patch is inputted into ADNET, it is firstly resized to 112*112*3, and then its deep CNN features are extracted from covn1 to fc6 and fc7 for action prediction and class prediction, respectively. The action and classification that have the max probabilities are selected, and the history c is also updated. The action selection strategy is trained by reinforcement learning.
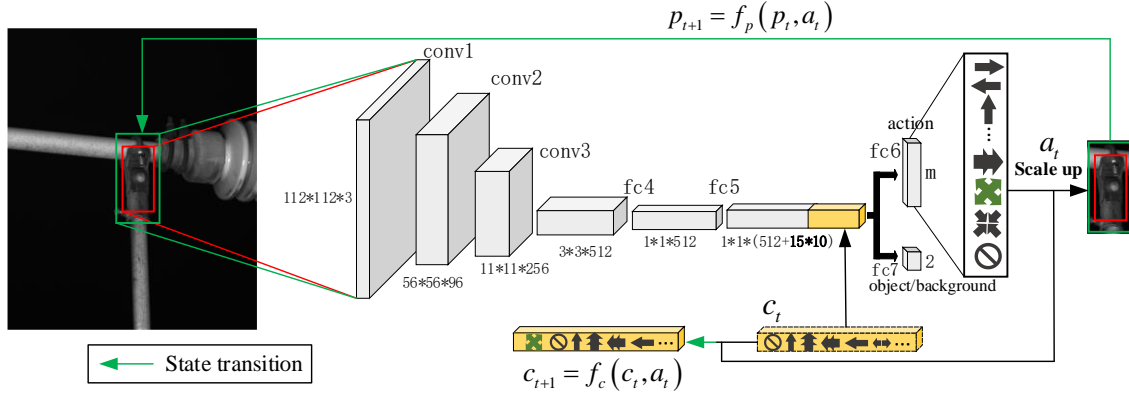
Fig. 3 The architecture of the proposed method.

III. REINFORCEMENT LEARNING FOR BS

A. MDP Formulation

The proposed localization refinement strategy follows a Markov Decision Process (MDP). This setting provides a formal framework to model an agent that makes a sequence of decisions. The MDP is defined by states $s \in S$, actions $a \in A$, state transition function $s' = f(s, a)$, and the reward $r(s, a)$. Here, we take the ADNET as an agent to find accurate box regions for BS component by taking sequential actions. Through formulating the localization refinement as the MDP, the action policy of ADNET can be optimized by reinforcement learning. The action, state, state transition function and reward are formulated as follows.

Action: Make the initial box fit the position and shape of BS, transformations of moving {*left, right, up, down*}, scale changing {*scale up, scale down*}, and shape changing {*fatter, taller*} are defined as possible actions. Especially, when the agent finds the optimum location, or the current localized box is the same as the previous box, a stop action is needed to finalize the box searching. Specifically, we define the action space A as shown in Fig. 4. The space A consists of 15 actions, and provides sufficient transform options for box changing. Rotating options are not considered in this paper.
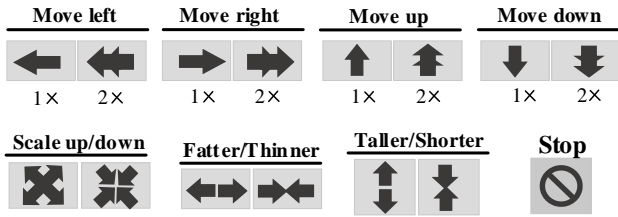
Fig. 4 The defined actions in our method.

State: As the localization refinement is a process of changing the geometry of box, using the information of what actions ADNET has done before is helping to predict better boxes [18, 19]. Thus, the image patch within box and the history actions are used to form the state s. For localization refinement in image I at step t, the state s_t is defined as a tuple (p_t, c_t), where $p_t \in R^{112*112*3}$ is the image within the current box and the $c_t \in R^{150}$ denotes the encoded vector of action history. The p_t can be formulated as,

$$p_t = \phi([\mathrm{x}_t, \mathrm{y}_t, \mathrm{w}_t, \mathrm{h}_t], \mathrm{I}) \qquad (1)$$

where (x_t, y_t) is the coordinate of center point of p_t in image I, w_t and h_t are the width and height of p_t respectively. The function ϕ crops p_t from image I and resizes it to the input size of ADNET. The c_t is a 150-dimensional vector, because we choose previous ten actions as history, and each action is encoded by 15 dimensions.

State transition function: When the ADNET selects an action a_t, the current s_t will transit to s_{t+1}. The state transition is performed by two functions $f_p(p_t, a_t)$ and $f_c(c_t, a_t)$, which are implemented on current image patch p_t and current action history c_t respectively. As for the f_p, the discrete amount of action transformations should be given. The discrete amounts of moving actions are given in (2). In (3), the discrete amounts of scale changing and shape changing actions are defined. As the initial box is not far away from the BS, we set factors α_1, α_2, α_3 and α_4 to 0.05 in our experiments.

$$\triangle x_t = \alpha_1 * w_t, \quad \triangle y_t = \alpha_2 * h_t \qquad (2)$$

$$\triangle w_t = \alpha_3 * w_t, \quad \triangle h_t = \alpha_4 * h_t \qquad (3)$$

As for the state transition f_c, it adds the current action a_t into action history c_t as the latest action, and removes the earliest action.

Reward: The reward can be regarded as feedback after taking an action. During the reinforcement learning training, if the selected action can make the state transition to a better state, then the agent will get a positive reward. Otherwise, a zero reward or negative reward will be returned. In this paper, the reward function $R_t(p_t, p_{t+1})$ is defined as follows.

$$R_t(p_t, p_{t+1}) = \begin{cases} +1, & if\ IoU(p_{t+1}, G) > IoU(p_t, G) \\ 0, & if\ IoU(p_{t+1}, G) = IoU(p_t, G) \\ -1, & if\ IoU(p_{t+1}, G) = IoU(p_t, G) \end{cases} \qquad (4)$$

where G is the ground-truth box of target BS, the $IoU(p_t, G)$ denotes overlap ratio of the current patch p_t and the ground truth G of the target BS with intersection-over-union criterion. In (4), only when the next patch gets closer to the ground-truth of BS than the case of current patch, the agent can obtain a positive reward.

B. Training Objective

The agent ADNET is trained by RL, whose goal is to learn a state-action policy that makes a sequence of action decisions. Before applying RL, we initialize the ADNET by utilizing the weight parameters trained by supervised learning

102

(SL), which has been proved useful for policy learning [13, 19].

In the SL stage, training samples p_i {i=1, 2, …, N} are generated by imposing Gaussian noise on the BS ground-truth G_i [x_i, y_i, w_i, h_i]. The action label $o_i^{(act)}$ and class label $o_i^{(cls)}$ are defined by (5) and (6), respectively.

$$o_i^{(act)} = \arg\max_a IoU\left(f_p(p_i,a),G_i\right), a \in A \quad (5)$$

$$o_i^{(cls)} = \begin{cases} 1, & if\ IoU(p_i,G) > 0.6 \\ 0, & otherwise \end{cases} \quad (6)$$

where A is the action space described in Section III, it has 15 actions. Then, the initial weight W_{SL}, {w_1, w_2, …, w_7}, is learned by minimizing the loss L_{RL}.

$$L_{SL} = \frac{1}{N}\sum_{i=1}^{N}[L\left(o_i^{(act)},\hat{o}_i^{(act)}\right) + L\left(o_i^{(cls)},\hat{o}_i^{(cls)}\right)] \quad (7)$$

The loss L_{SL} includes action prediction loss and classification (object/background) loss. N is batch size of training samples and L denotes the cross-entropy loss. The $\hat{o}_i^{(act)}$ and $\hat{o}_i^{(cls)}$ are the predicted action and predicted class for sample i, respectively. Note that the history action vector is not used in this stage.

In the RL training stage, the W_{RL}, {w_1, w_2, …, w_6}, is initialized by W_{SL}. The fc7 is ignored because only action is concerned in RL. For an image frame l, an initial box will take T actions that are successively predicted by ADNET. In each step t, new features are drawn from the image, allowing the RL algorithm to adapt to new information. The action $a_{t,l}$ is selected by

$$a_{t,l} = \arg\max p\left(a \mid s_{t,l};W_{RL}\right) \quad (8)$$

where $t = 1, 2, …, T-1, T$.

After action $a_{t,l}$ has been taken, the ADNET gets a reward $R_{t,l}$ according to (4). Meanwhile, a history of what actions ADNET has done before is also recorded. Then the history action vector c_t is updated by adding $a_{t,l}$ and removing the earliest action.

Finally, W_{RL} is updated using stochastic gradient ascent [23] to maximize the accumulated rewards of the training samples as follows.

$$\Delta W_{RL} \propto \sum_{l}^{L}\sum_{t}^{T_i}\frac{\partial \log p\left(a \mid s_{t,l};W_{RL}\right)}{\partial W_{RL}}R_{t,l} \quad (9)$$

where L is the size of image patches that used at one iteration.

IV. EXPERIMENTAL RESULTS

A. Dataset and Training Setting

The dataset in our system is collected from the Changsha-Zhuzhou high-speed rail line in China. It has 1596 catenary BS images that cropped from global catenary images. Each BS image is annotated with a tight box of ground-truth position. As BSs component have different sizes in global images. The width (or height) of each cropped BS image is 2.5 times of the truth width (or height) of the BS component, which makes the ADNET have proper regions for researching. We use 1020 images for training and 576 images for testing.

To train ADNET, we set the learning rate to 0.0001 for conv1-conv3 and 0.001 for fully-connected layers (fc4-fc7), momentum to 0.9, weight decay to 0.0005, and mini-batch size to 64. The epoch numbers of training iteration are set to be 200 for SL and 300 for RL, respectively. The experimental environment of reinforcement learning is as follows: Linux Ubuntu 14.04, MATLAB 2017a, CUDA 8.0 and NVIDIA GTX1080Ti GPU with 11 GB memory.

B. Experiment and Analysis

We evaluated our method on the built dataset. As the initial localized boxes can be distributed in any position around the BS component. Therefore, we apply a Gaussian function on the ground-truth BS position to produce the initial boxes. In the testing, the agent ADNET takes T consecutive actions (steps) in each image to refine the initial boxes. Here, T is set to be 20. Some selected localization results are shown in Figs.5-8. Performances of the proposed method over the entire dataset are summarized in Table 1. Detailed experiment analyses are as follows.

1) Visualization and analysis of RL-based refinement.

The dynamic processes of localizations by the agent are displayed in Figs.5-7, where the white box is the initial box and its color gets greener gradually after taking an action until reaching the final pink box. We divide these processes into three types, namely *Type A*, *Type B* and *Type C*, which are shown in Figs.5-7 respectively.
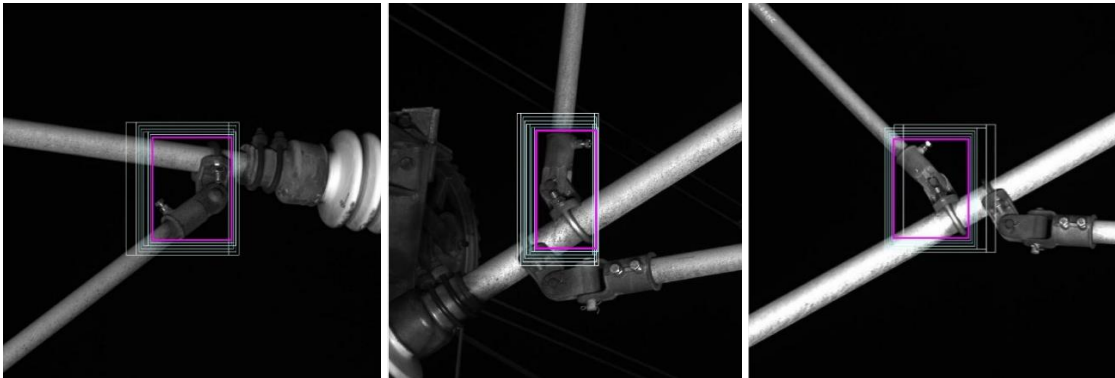
Fig. 5 Localization refined by RL agent for *Type A* cases. **Left**: case A1. **Middle:** case A2. **Right:** case A3.
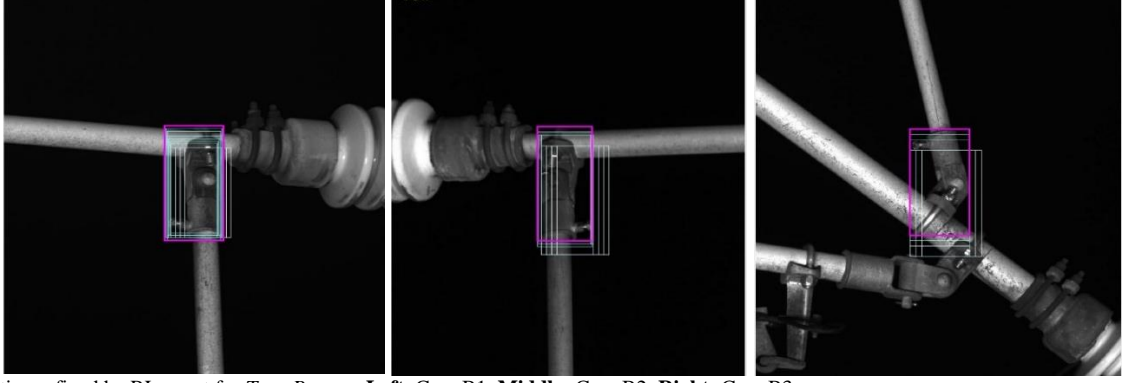
Fig. 6 Localization refined by RL agent for *Type B* cases. **Left**: Case B1. **Middle:** Case B2. **Right:** Case B3.
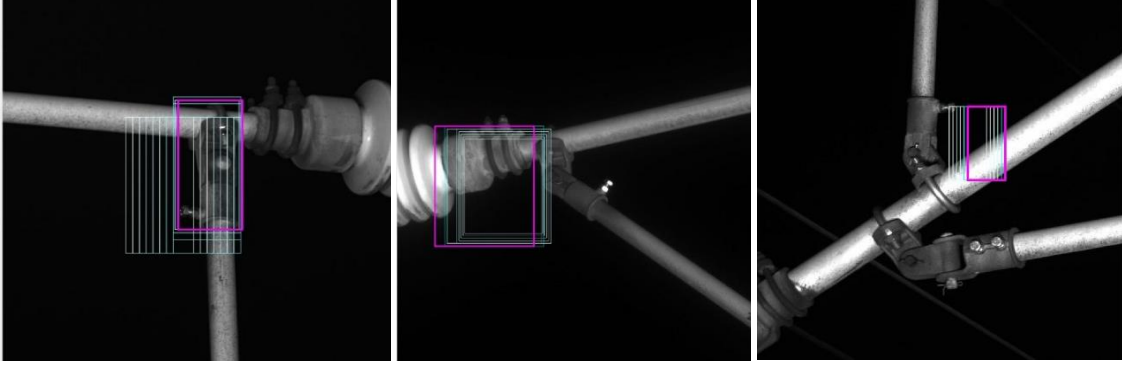
Fig. 7 Localization refined by RL agent for *Type C* cases. **Left**: case C1. **Middle:** case C2. **Right:** case C3.
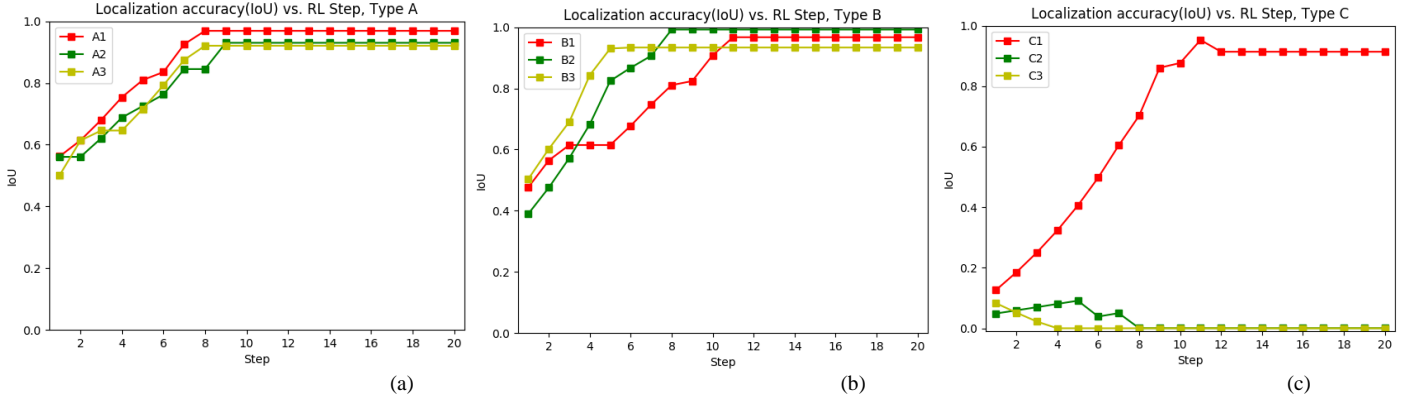
Fig. 8 Localization accuracy (IoU) changes for three different types of cases in the testing. (a) IoU changes in *Type A* cases. (b) IoU changes in *Type B* cases. (c) IoU changes in *Type C* cases

In Fig. 5, when the *Type A* initial box is larger than the BS component, and it may contain unnecessary information from other components, such as case *A2* and case *A3*. The agent can adjust these initial boxes closer to BS components mainly by taking actions including *moving*, *scale down*, *shorter* and *thinner*. Fig. 8(a) shows the localization accuracy (IoU) changes of *Type A* cases when the step grows. The IoUs are getting larger even the initial boxes are with unnecessary information. In Fig. 6, the initial boxes of *Type B* cases are partly overlapped with the BSs. Particularly, the initial box of case *B3* contains not only the incomplete BS, but also unnecessary information from near components. However, the RL-learned agent can still move these initial boxes closer to the BSs. The accuracy changes of each case are shown in Fig. 8(b). Both *Type A* and *Type B* are cases that can be correctly refined by the proposed method. As for the speed of refinements, although we set the total steps for each image to

20, most of the successful cases take less than 12 steps to reach the final locations, as shown in Fig. 8 (a) and (b).

As we use the Gaussian function to randomly produce the initial boxes, some initial boxes have fewer overlaps with BS component, as the *Type C* cases showed in Fig. 7. The initial (first step) IoUs of *Type C* cases can be observed in Fig. 8(c). They are less than 0.15 at the beginning. In the experiment, many *Type C* cases led to failed BS localizations and the boxes are moved further away from the BS, as shown by the case *C2* and case *C3* in Fig. 7. However, there are still some cases like case *C1* can successfully adjust boxes closer to the BS component position. We conjecture that case *C2* and case *C3* are failed because very few features of BSs are contained within their initial boxes, while case *C1* still contains some useful features of BS, which can be observed from the IoUs at the first step in Fig. 8(c).

2) Quantitative analysis of RL-based refinement result.

The localization accuracy metric for the overall dataset is the widely used *Recall* [24]. The **Recall**$_{TIoU}$ **(IoU>0.5)** means the proportion of having a IoU larger than the threshold T_{IoU}. The speed metric adopted is FPS (frame/second). Quantitaive localization performances of the proposed method is shown in Table 1.

Table 1. Localization accuracy improvement by RL

Method	Recall$_{0.5}$ (IoU>0.5)	Recall$_{0.8}$ (IoU>0.8)	Recall$_{0.9}$ (IoU>0.9)	Improved proportion	FPS
Gaussian Initial	96.4%	19.9%	3.3%	--	--
Gaussian Initial with RL (ours)	92.0%	**79.2%**	**65.1%**	**81.8% (471/576)**	6.13

Table 1 shows that the Recall$_{0.5}$ is slightly decreased compared with the initial value, because few boxes are moving away from the BS positions. However, the Recall$_{0.8}$ is increased from 19.9% to 79.2%, and the Recall$_{0.9}$ is increased from 3.3% to 65.1%. Overall, among 576 test images, 471 images' IoUs become larger, which means 81.8% test images get better localizations. It indicates that the proposed RL method can adjust most of boxes closer to the BS positions and improve the localization accuracy. Besides, the RL agent takes only 0.163s (1/6.13) for each test image, which consumes very little time in applications.

V. CONCLUSION

This paper proposes a novel approach for improving the localization accuracy of BS (Brace Sleeve) in railway catenary systems. Differing from the existing localization strategies in railway that localize objects following a single structured prediction model, the proposed method adopts a dynamic searching strategy. We investigate one method that adopts RL to train an agent to generate an improved bounding box for BS localizations through a sequence of defined action transformations. Experimental results using real-life inspection images show that the proposed method can adequately and effectively refine the localization from a coarse-localized input. Nevertheless, there are still some further improvements to be conducted:

(1) As the case *C2* and case *C3* shown in Fig. 7 and the Recall$_{0.5}$ comparison shown in Table 1, there are some failed cases with the bounding boxes moving away from the BSs. This issue should be further researched and improved.

(2) When a BS is not localized by our agent, exploring and implementing a new searching strategy in the failed image will also reduce the number of failed cases.

(3) Except for the policy gradient-based reinforcement learning method [23] used in this paper. Elements from applications of reinforcement learning in other fields, such as robotics and control [25, 26], can also be considered for the dynamic monitoring of catenary systems.

REFERENCES

[1] S. Gao, Z. Liu, L. Yu., "Detection and monitoring system of the pantograph-catenary in high-speed railway (6C)," in 7th *International Conference on PESA*, Hongkong, pp. 779-788, 2017.

[2] Y. Han, Z. Liu, X. Geng, and J. P. Zhong, "Fracture detection of ear pieces in catenary support devices of high-speed railway based on HOG eatures and two-dimensional Gabor transform," *J. China Railway Soc.*, vol. 39, no. 2, pp. 52–57, 2017.

[3] J. Zhong, Z. Liu, G. Zhang, and Z. Han, "Condition detection of swivel clevis pins in overhead contact system of high-speed railway," *J. China Railway Soc.*, vol. 39, no. 6, pp. 65–71, Jun. 2017.

[4] H. Fan, P. Cosman, Y. Hou, et al, "High-speed railway fastener detection based on a line local binary pattern," *IEEE Signal Processing Letters.*, vol. 25, no. 5, pp. 788-792, 2018.

[5] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, Jun. 2015.

[6] Z. Cai and N. Vasconcelos, "Cascade R-CNN: delving into high quality object detection," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 6154-6162, 2018.

[7] Z. Liu, L. Wang, C. Li, et al., "A high-precision loose strands diagnosis approach for isoelectric line in high-speed railway," *IEEE Transactions on Industrial Informatics,* 10.1109/TII.2017.2774242.

[8] Z. Liu, Y. Lyu, L. Wang, et al., "Detection approach based on an improved faster RCNN for brace sleeve screws in high-speed railways," *IEEE Transactions on Instrumentation & Measurement*, vol. 69, no. 7, pp. 4395-4403, 2020.

[9] G. Q. Kang, S. B. Gao, L. Yu, and D. Zhang, "Deep architecture for high-speed railway insulator defect detection: denoising autoencoder with multitask learning," *IEEE Transactions on Instrumentation and Measurement*, DOI: 10.1109/TIM.2018.2868490.

[10] Z. Liu, K. Liu, J. P. Zhong, Z. Han and W Zhang, "A high-precision positioning approach for catenary support components with multi-scale difference," *IEEE Transactions on Instrumentation & Measurement*, vol. 69, no. 3, pp. 700-711, 2020.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779-788.

[12] J. Chen, Z. Liu, H. Wang, et al., "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network," *IEEE Transactions on Instrumentation and Measurement*, 10.1109/TIM.2017.2775345.

[13] D. Silver, A. Huang, C. J. Maddison, et al., "Mastering the game of go with deep neural networks and tree search," *Nature*, 529(7587): 484–489, 2016.

[14] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning,". *CoRR*, abs/1509.06461, 2015.

[15] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," *In ICML*, 2014.

[16] F. Ruelens, B.J. Claessens, S. Quaiyum, et al., "Reinforcement learning applied to an electric water heater: From theory to practice," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3792-3800, 2018

[17] V. Mnih, K. Kavukcuoglu, D. Silver, et al., Playing atari with deep reinforcement learning. *arXiv preprint* arXiv:1312.5602, 2013.

[18] J. C. Caicedo and S. Lazebnik, "Active object localization with deep reinforcement learning," *IEEE International Conference on Computer Vision*, pp. 2488–2496, 2015.

[19] S. Yun, J. Choi, Y. Yoo, et al., "Action-Decision networks for visual tracking with deep reinforcement learning," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2711-2720, 2017.

[20] M. Bellver, X. Giro-I-Nieto, F. Marques, et al., "Hierarchical object detection with deep reinforcement learning," *arXiv preprint arXiv:1611.03718*, 2016.

[21] S. Liu, D. Huang and Y. Wang, "Pay attention to them: deep reinforcement learning-based cascade object detection," *IEEE Transactions on Neural Networks and Learning Systems*, PP (99):1-13. 2019.

[22] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv: 1409.1556*, 2014.

[23] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine Learning*, 8(3-4):229–256, 1992.

[24] B. Michael, and G. Fredric, "The relationship between recall and precision," *Journal of the American Society for Information Science*, vol. 45, no. 1, pp. 12-19, 1994.

[25] Y. Pane, S. Nageshrao, J. Kober, et al., "Reinforcement learning based compensation methods for robot manipulators," *Engineering Applications of Artificial Intelligence*, 78:236–247, 2019.

[26] T. de Bruin, J. Kober, K. Tuyls, et al., "Integrating state representation learning into deep reinforcement learning," *IEEE Robotics and Automation Letters*, 3(3):1394–1401, 2018.

An Automatic Defect Detection Method for Catenary Bracing Wire Components Using Deep Convolutional Neural Networks and Image Processing

Wenqiang Liu*
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
Liuwq_2009@126.com

Dan Wang
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
17454845@qq.com

Cheng Yang
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
yangc@my.swjtu.edu.cn

Yuyang Li
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
824285241@qq.com

Hui Wang
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
1939563058@qq.com

Zhigang Liu
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
liuzg_cd@126.com

Abstract—**Bracing wire components, including messenger wire bases, bracing wire hooks, and bracing wires, are essential to make certain the stability of high-speed catenary support devices. The core process of the proposed detection method includes two stages, the first is that brace wire components are localized first, and the second stage is that the defects of the components are detected. In this article, a new defect detection method for catenary bracing wire components is proposed. First, Faster R-CNN, one of the advanced deep convolutional neural networks, is introduced to localize messenger wire bases and bracing wire hooks, accurately. And then, the regions of bracing wires are extracted according to the structure information among bracing wire components and bracing wires are localized by the Hough Transformation. Next, the installation defects of messenger wire bases are detected with image processing methods, and the looseness defects of bracing wires are detected by combining with the results detected by the Hough Transform. Experiment results prove that the proposed scheme can localize and identify the defects of bracing wire components, accurately.**

Keywords—High-speed railway, catenary bracing wire components, defect detection, deep convolutional neural networks, image processing.

I. INTRODUCTION

With more and more prominent advantages of high-speed railway, countries around the world have built a large number of railways. The railway foundation mainly consists of two parts: the catenary system and the rail system. The catenary system is primarily responsible for the power supply of high-speed locomotives, and it contains a large number of support components, such as insulators, steady arm bases, bracing wire hooks, etc. (as shown in Figure1). These support components are used to withstand mechanical loads, electrical insulation, etc. However, due to the high-speed movement of the locomotive and the influence of the external environment, the catenary support components may have the loss and looseness of parts, cracks, etc., which will pose a significant threat to the safe operation of high-speed trains. In this context, fast and accurate target localization and defect detection methods are particularly important. The

*Wenqiang Liu is the corresponding author. (e-mail: Liuwq_2009@126.com)
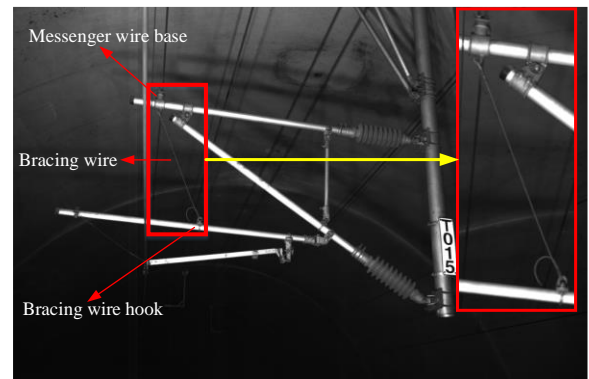
Fig. 1. The bracing wire components of catenary support devices, including messenger wire bases, brace wires, and bracing wire hook.

speeded-up robust features (SURF) was proposed to extract and match the characteristics of the insulator in the image. And then, the state of the insulator was identified and located according to the distribution rule of the minimum gray value [1]. In [2], the features of scale-invariant feature transform (SIFT) of pantograph were extracted to locate the pantograph, evaluat-ting the reliability of railway catenary power supply by measuring the stagger between pantographs and contact wires.

In [3], a detection and identification method for high-speed railway rod insulators using Chan-Vese (CV) model and contour transformation (CT) was proposed. First, the directionality and anisotropy features of catenary insulator images were detected using CT method in the frequency domain. Then, the insulator boundaries were identified by the CV model. In [4], a detection and localization method for catenary insulators was proposed. First, a sub-image containing strip devices was extracted from the image. Then, by combing the latent support vector machine (SVM) and deformable part model, catenary insulators were identified and detected from these strip sub-images. A failure detection method for catenary earpieces was proposed in [5], that adopted histogram of oriented gradients (HOG) to obtain the features of rotary double-ears and combined with SVM to recognize them. And 2D Gabor wavelet transform was used to achieve a reliable diagnosis of the fracture failure of earpieces. With the improvement of computer computing
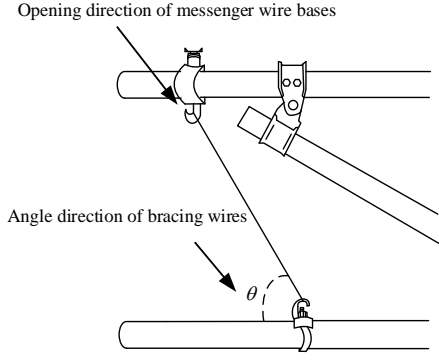
Fig. 2. Schematic diagram of the normal state of bracing wire components.
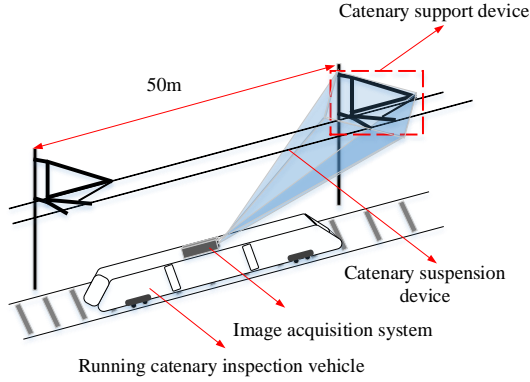
Fig. 3. The data collection system of catenary inspection vehicle

power and information collection ability, many deep learning-based algorithms are proposed. In [6], steady arm bases were detected using the region-based convolutional neural networks (R-CNN) first. And then, steady arms were detected with Hough transform based on the installation relationship, and finally the slope of the steady arm was calculated. In [7], the crack defect of messenger wire supporters was detected based on Faster R-CNN. In [8], a method for diagnosing the looseness of isoelectric lines based on deep learning was adopted to. In [9], an object detection network combined with depth-wise separable convolution model was adopted to achieve the detection of the dropper fault was detected. First, they used the dropper progressive location network (DPLN) to detect the dropper, then, a fault recognition network (DFRN) for the dropper was utilized to identify dropper fault types. In [10], a multi-algorithm-fused image processing method was proposed to achieve the defect detection that includes, dropper-strands loosened, micro deformation, dropper-strands broken, and so on. Although these methods have achieved some good results, the defect detection for bracing wire components has rarely been reported.

In Figure 2, the schematic diagram of the normal state of bracing wire components is shown. Under normal conditions,

bracing wires present a straight line due to the force, and the opening direction of the messenger wire base installation should be consistent with the direction of the angle θ of bracing wires. However, due to the carelessness of the installation staff, messenger wire bases and bracing wire hooks are often easily reversed. Besides, due to the long-term operation of catenary support devices, bracing wires are prone to looseness. Therefore, a detection method for bracing wire defects based on the deep convolutional neural network (DCNN) is proposed in this paper.

The arrangement of this paper is introduced as follows: In Section II, the system overview is introduced first. Next, the proposed method is expressed in Section III. Furthermore, in Section IV, experiment results are analyzed and discussed. Lastly, some conclusions are summarized in Section V.

II. SYSTEM OVERVIEW

The catenary support component images are collected by the high-speed cameras mounted on the roof of the running catenary inspection vehicle shown in Figure 3, continuously. And the acquired image resolution is 6600×4400. Before the catenary images are put into the detection network in the first stage, the images are resized into 990×660 to improve the detection efficiency of the detection network. The pipeline of the defect detection module mainly consists of two main image-processing stages, component localization, and defect detection, as shown in Figure 4. The detection system overview is described as follows.

A. Component Localization

First, messenger wire bases and bracing wire hooks are extracted and localized with the DCNN in the first stage. Next, the region of bracing wires is extracted according to the structure information among messenger wires, bracing wires, and bracing wire hook, and then the bracing wire is localized with the Hough Transformation in the extracted region.

B. Defect Detection

First, the installation defects of messenger wire bases are detected by the proposed image processing method. Next, the looseness defects of bracing wires are detected using the detected results with Hough Transform.

III. METHODOLOGY

In this research, a new defect detection method is presented for catenary bracing wire components. And to achieve the goal, the methodology of the proposed method is introduced as follows.

A. Component Localization Based on Faster R-CNN

In [11], the performance of the component localization of catenary support devices using the advanced DCNNs were
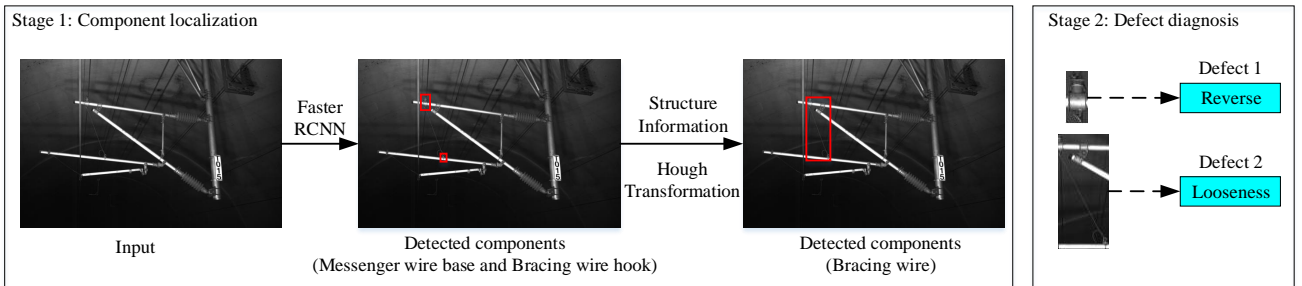
Fig. 4. Pipeline of the proposed detection system.
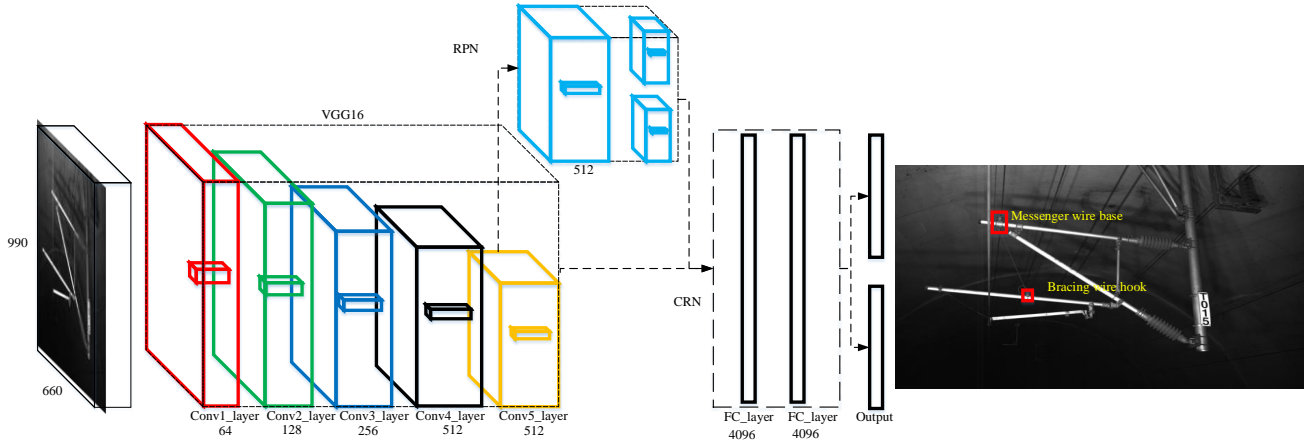
Fig. 5. Localization of messenger wire bases and bracing wire hooks with the Faster R-CNN.

(a)   (b)
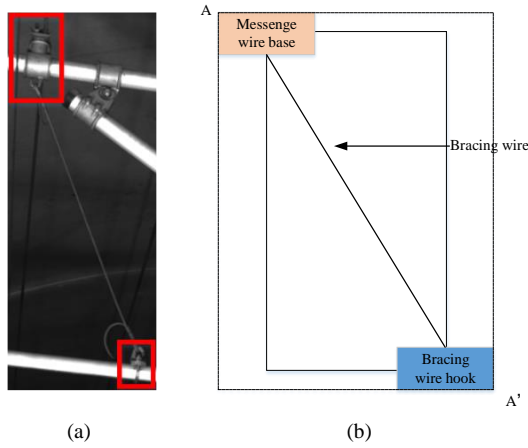
Fig. 6. The diagram of region extraction and localization of bracing wires based on the structure information of bracing wire components. (a) is a real image of bracing wire components, (b) shows the structure information of bracing wire components with block diagrams.

(a)   (b)

(c)   (d)

Fig. 7. The process of image rotation. (a) is an original image, (b) shows the binarized image, (c) denotes the detected result of horizontal cantilever with Hough Transformation, (d) is the rotated images.
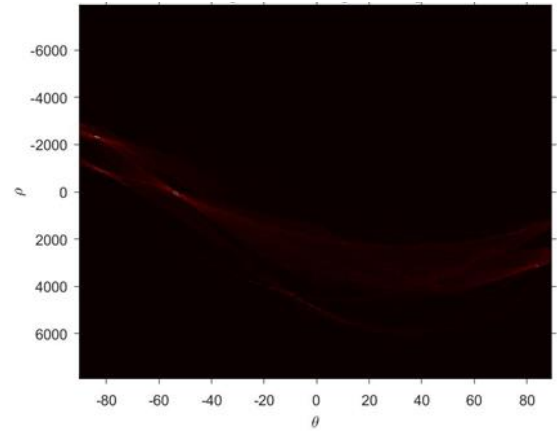
Fig. 8. Hough Transform for the horizontal cantilever.

compared and evaluated, and the results shown that the Faster-RCNN had a better result. Therefore, Faster R-CNN is chosen to locate components in this paper. The pipeline of Faster R-CNN in [12] includes three modules in total, as shown in Figure 5. In the first module, convolutional neural networks are used to extract the features of objects. In the second module, region proposal networks are adopted to predict the regions of interest (RoIs) of objects. And fully connected networks are utilized to classify objects and regress their locations in the third module. In this research, the VGG16 in [13] is adopted as the feature extraction network. And the localization process of bracing wire components is as follows.

(1) Localization of messenger wire bases and bracing wire hooks

According to the feature information of messenger wire bases and bracing wire hooks, the object detection network Faster R-CNN is trained with catenary images, which can accurately detect the messenger wire base and bracing wire hook components, as shown in Figure 5.

(2) Localization of bracing wires

According to the structure information of bracing wire components, the bracing wires must be contained within the AA' region. And then, the bracing wires are localized with the Hough Transformation in the extracted region, as shown in Figure 6.

B. Defect Detection Using Image Processing

As above introduced, the defects of bracing wire components mainly include two types, the installation reverse of messenger wire bases and the looseness of bracing wires. To accurately detect these two defects, some image processing methods are adopted as follows.

(1) Looseness defect of bracing wires

The detection method for the looseness defect includes the following steps.

Step 1) Image rotation

108

To facilitate the analysis of the opening direction, the image should be rotated to horizontal according to the angle of the horizontal cantilever, as shown in Figure 7. First, the image is binarized, as shown in Figure 7(b). Next, the edge of the binarized image is extracted by the Candy [14], and the horizontal cantilever is detected with the Hough Transform, as shown in Figure 8. And according to the detection result, messenger wire bases are rotated to horizontal. The Hough Transform formula is as follows.

$$\rho = x\cos\theta + y\sin\theta \tag{1}$$

(a)   (b)
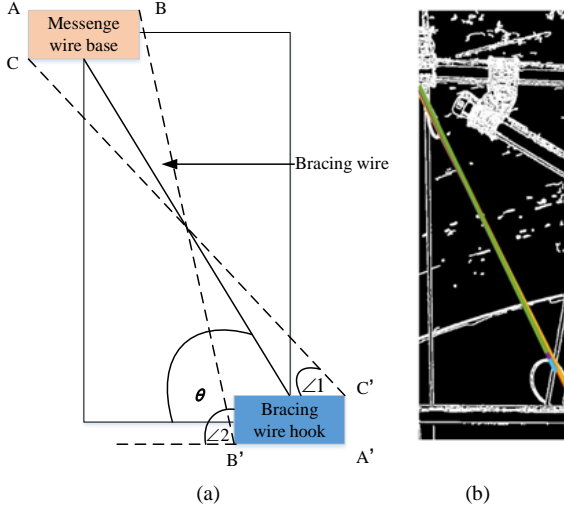
Fig. 9. The schematic diagram of the bracing wire detection. (a) is a parameter illustration of the bracing wire detection, (b) shows a detected bracing wire image.
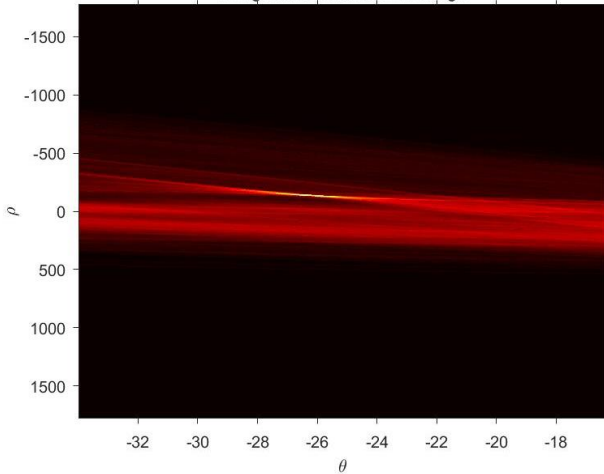
Fig. 10. Hough Transform for bracing wires.
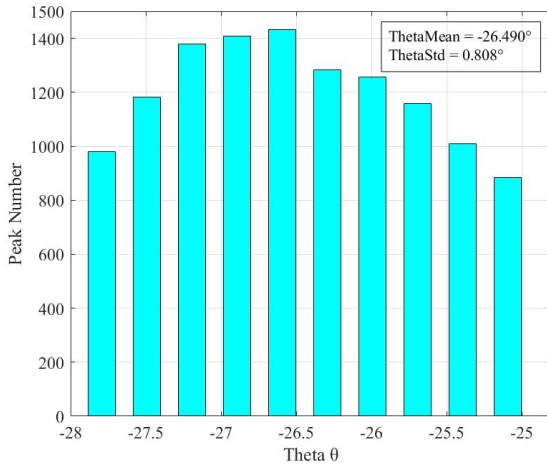
Fig. 11. Peak distribution of theta θ.

Step 2) Bracing wire detection with Hough Transform

To accelerate the line detection and improve the detection accuracy, a fast and accurate detection scheme with Hough Transform is proposed. As shown in Figure 9 (a), according to the structure information of bracing wire components, the angle θ of bracing wires must be in the range of the angle $\angle 1$ and angle $\angle 2$. Therefore, based on the characteristic, the line detection with Hough Transform can eliminate straight lines from other angles, which ensures the detection accuracy and efficiency. The result is shown in Figure 9 (b).

Step 3) Defect detection

In Figure 10, the Hough Transform results between the angle $\angle 1$ and angle $\angle 2$ are shown. And the peak distribution of θ ranked in the top ten is plotted in Figure 11. In theory, the straighter the line, the smaller the angle fluctuations, and the number of the peak should be within a range. Therefore, the standard deviation *thetaStd* of θ and is used as the indicator to detect the looseness defect. As shown in Figure 11, If *thetaStd* is less than 1, then the line is determined to be normal. Otherwise, the line is determined to be the looseness defect. The formulas are as follows.

$$thetaStd = \sqrt{\frac{\sum_{i=1}^{n}(\theta_i - \overline{\theta})^2}{n}} \tag{2}$$

$$state = \begin{cases} normal & thetaStd \leq 1 \\ failure & thetaStd > 1 \end{cases} \tag{3}$$

(2) Installation reverse defect of messenger wire bases

The detection method for the installation reverse defect includes the following steps.

Step 1) Messenger wire base extraction with image preprocessing

To facilitate the defect detection, the image needs to be preprocessed, as shown in Figure 12, and Figure 13.

First, since the detection is susceptible to background interference in the tunnel environment, the background should be eliminated first, as shown in Figure 12.

a) The grayscale histogram $Hist(k)$ of the image is counted by Equation 4, as shown in Figure 12(b);

b) The median *Median* of the image is calculated by Equation 5. And through setting the pixel values below the median to 0, the image with the background eliminated is obtained in Figure 12(c). And its new grayscale histogram is counted in Figure 12(d).

Next, the image is further processed through binarization, dilation, erosion, and filling, as shown in Figure 13.

$$Hist(k) = \sum_{k=1}^{L}\sum_{i=1}^{M}\sum_{j=1}^{N} Hist(k)+1 \ \ if \ \ (k-1)\leq I_{(i,j)}\leq k \tag{4}$$

$$Median = \frac{1}{2}(Rank(I_{ij})_{\frac{M*N}{2}} + Rank(I_{ij})_{\frac{M*N}{2}+1}) \tag{5}$$

where $I_{(i,j)}$ presents the grayscale of the image pixel, L presents the grayscale level, (M,N) denote the size of the image, $Rank(\cdot)$ indicates a sort function by the grayscale value of the image.
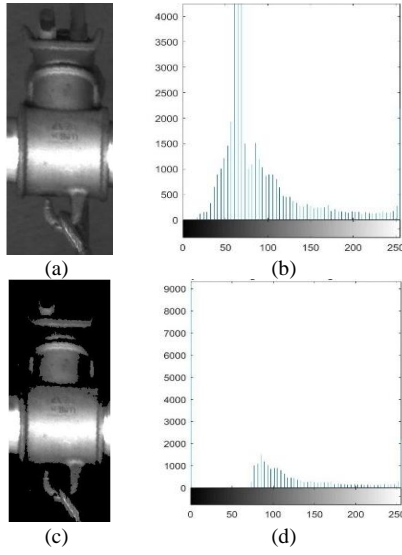
Fig. 12. The process of background elimination. (a) is an original image, (b) indicates the grayscale histogram, (c) presents the image with the background eliminated, and (d) shows the grayscale histogram of the image with the background eliminated.
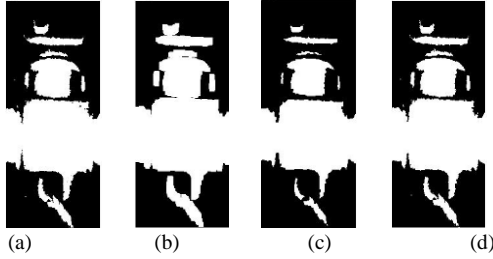
Fig. 13. Image opening and closing operation. (a) binarization, (b) dilation, (c) erosion, (d) filling.

(a)

(b)
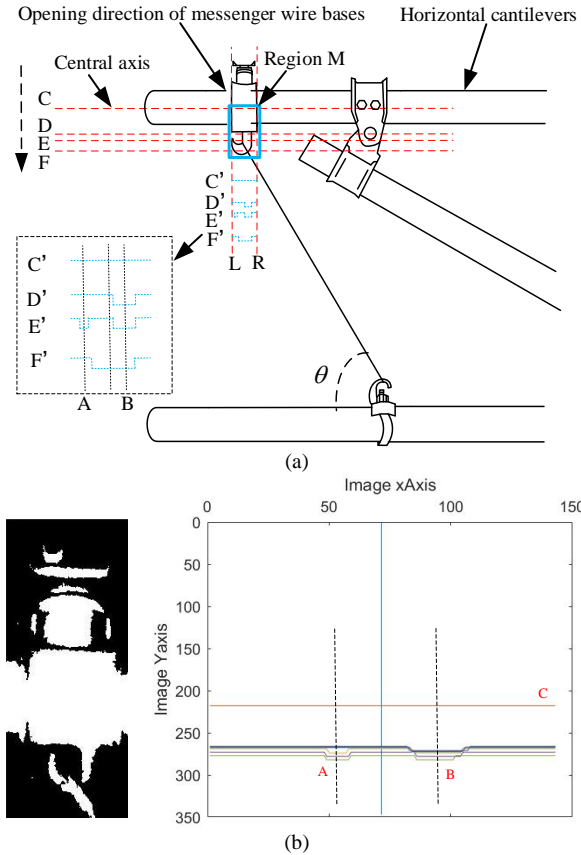
Fig. 14. The opening direction detection. (a) is the schematic of the opening direction detection, (b) presents a result example of the opening direction detection.

Step 2) Defect detection

As shown in Figure 14 (a), the schematic of the opening direction detection is shown. And through scanning according to the horizontal direction to generate pulse signals, the defect of the installation reverse is detected. First, as shown in Figure 14 (a), messenger wire bases are installed on the horizontal cantilever, so the processed image in Figure 13 (d) begins to scan from the central axis A of horizontal cantilever down. When the opening direction is to the left, as shown in Figure 14(b), one pulse signal appears at B at first, and then continuing to scan, two pulse signals appear at A and B. And if the opening direction is to the right, then the direction of the detected signal is reversed. Therefore, according to the characteristic, the opening direction can be determined.

Once the opening direction of the messenger wire bases and the angle θ of bracing wires are determined, the installation defect can be detected. If they are consistent in their direction, then the component is normal. Otherwise, the component is the installation reverse.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

To valid the efficiency of the proposed method, some relevant experiments are implemented below. The experiment environment is based on the Linux 17.10 platform and the program development framework adopts the TensorFlow [18]. The system configuration is the 32 GB RAM, GeForce GTX 1080Ti, and the intel core i7-8700 CPU @3.70GHz×12 processor.

A. Data Set

As shown in Figure 3, the catenary image data set is collected from the running catenary inspector vehicle in the field. The total amount of the catenary image data set is 4644, the training and validation data are 2275 and 975, respectively, and the test data are 1394. The training parameters of Faster R-CNN are initialized as follows: the VGG16 network is initialized by the pre-trained parameters, weight decay and momentum are initialized into 0.0001 and 0.9, respectively. The iteration time initialized into 10,000, and the learning rate is initialized into 0.0001. And after 10000 times, the learning rate is adjusted to 0.00001, continuing to train 5000 times.

B. Evaluation Indexes

The evaluation metrics introduce precision (*P*), recall (*R*), average precision (*AP*), and the precision-recall (*PR*) curve.

$$P = \frac{TP}{TP + FP} \times 100\% \qquad (6)$$

$$R = \frac{TP}{TP + FN} \times 100\% \qquad (7)$$

$$AP = \int_0^1 P(R)\, \mathrm{d}R \qquad (8)$$

where *TP* presents true positive, *FP* denotes false positive, *FN* indicates false negative, and *TN* means true negative.

TABLE I. DETECTION FOR DIFFERENT DEFECTS

Components	Total samples	Detected samples
Messenger wire bases	30	27
Bracing wire	26	24
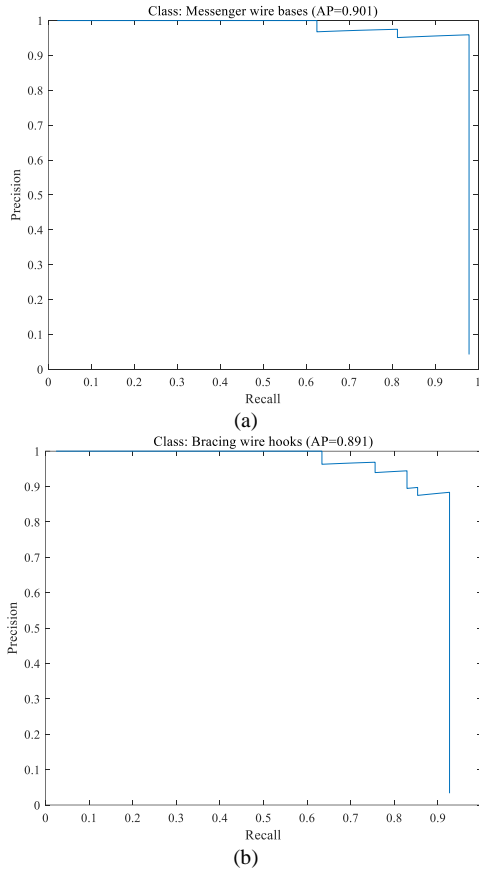
Fig. 15. The *PR* curves. (a) is the *PR* curve of messenger wire bases, (b) presents the *PR* curve of bracing wire hooks.
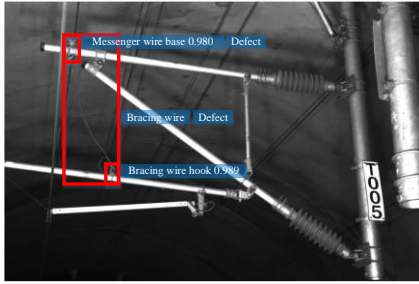
Fig. 16. Detection example.

C. Component Localization and Defect Detection

To analyze the performance of the proposed method, some experiments are performed as follows. First, the *PR* curves are drawn in Figure 15, and the APs of messenger wire bases and bracing wire hooks are 0.901 and 0.891, respectively. And through the analysis of three indexes, the proposed method can accurately localize the catenary components. And the detection results for different defects are described in Tab. I. And in Figure 16, a detection example is shown. The proposed method has achieved the goal of the defect detection for bracing wire components, but it still has some disadvantages under some special cases.

1) If only one of the messenger wire bases and bracing wire hooks is localized, the bracing wires will not be localized. And this would cause the defects of bracing wire components not to be detected.

2) In a picture, there may be multiple messenger wire bases and bracing wire hooks, they cannot match each other.

And this would lead the bracing wires to be wrongly localized, and the defects of bracing wire components not to be detected.

3) Because the light supplement of the image acquisition system may not fully fill the bracing wire, and the light is too dark, making it difficult to extract the bracing wire.

V. CONCLUSIONS

This paper presented a novel defect detection method for bracing wire components. The DCNN Faster R-CNN is adopted to localize messenger wire bases and bracing wire hooks. And for two types of defects, two defect detection methods based on image processing are presented, and corresponding criteria are given. The experimental results can prove the proposed method is capable of detecting the defects and has a good application prospect.

REFERENCES

[1] H. Yang, Z. Liu, Y. Han, and Z. Han. (2013). Defective Condition Detection of Insulators in Electrified Railway Based on Feature Matching of Speeded-up Robust Features. Power System Technology, 37(8), 2297-2302.

[2] C. Cho, and H. Ko. (2015). Video-Based Dynamic Stagger Measurement of Railway Overhead Power Lines Using Rotation-Invariant Feature Matching. IEEE Transactions on Intelligent Transportation Systems, 16(3), 1294-1304.

[3] G. Zhang, Z. Liu, and Y. Han. (2016). Automatic recognition for catenary insulators of high-speed railway based on contourlet transform and Chan–Vese model. Optik - International Journal for Light and Electron Optics, 127(1), 215-221.

[4] Y. Han, Z. Liu, D. Lee, G. Zhang, and M. Deng. (2016). High-speed railway rod-insulator detection using segment clustering and deformable part models. In Image Processing (ICIP), 2016 IEEE International Conference on, 3852-3856, IEEE.

[5] Ye. Han, Z. Liu, X. Geng, and J. Zhong. (2017). Fracture Detection of Ear Pieces in Catenary Support Devices of High-speed Railway Based on HOG Features and Two-Dimensional Gabor Transform. Journal of the China Railway Society, 39(2), 52-57.

[6] D. Chen, W. Zhang, and Y. Yang. (2017). Detection and recognition of high-speed railway catenary locator based on Deep Learning. Journal of University of Science and Technology of China, 47(4), 320-327.

[7] K. Liu, Z. Liu, and J. Chen. (2017). "Crack detection of messenger wire supporter in catenary support devices of high-speed railway based on Faster R-CNN," China Railway Society.

[8] Z. Liu, L. Wang, C. Li, and Z. Han. (2017). A High-Precision Loose Strands Diagnosis Approach for Isoelectric Line in High-Speed Railway. IEEE Transactions on Industrial Informatics,14(3), 1067-1077.

[9] S. Liu, L. Yu, and D. Zhang. (2019). An Efficient Method for High-Speed Railway Dropper Fault Detection Based on Depthwise Separable Convolution. IEEE Access, 7, 135678-135688.

[10] P. Tan, X. Li, Z. Wu, J. Ding, and J. Ma. (2019). Multialgorithm Fusion Image Processing for High Speed Railway Dropper Failure-Defect Detection. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 1-13.

[11] W. Liu *et al.*, (2018). "Multi-objective performance evaluation of the detection of catenary support components using DCNNs," *IFAC-PapersOnLine,* 51(9), 98-105.

[12] S. Ren, K. He, R. Girshick, and J. Sun. (2015). "Faster R-CNN: towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis & Machine Intelligence, volume (39), 1137-1149.

[13] K. Simonyan and A. Zisserman. (2015). Very deep convolutional networks for large-scale image recognition. Int. Conf. on Learning Representations, San Diego, CA.

[14] J. Canny. (1986). A Computational Approach To Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6):679-698.

Pantograph Vibration Interference Signal Recognition Based on SVM Classification

Yuming Ding
China Academy of Railway Sciences
Beijing, China
dingyuming@rails.cn

Jun Zhao
Railway Infrastructure Inspection Center
China Academy of Railway Sciences Corporation Limited
Beijing, China
zhaojun@rails.cn

Jinzhao Liu*
Infrastructure Inspection Research Institute
China Academy of Railway Sciences Corporation Limited
Beijing, China
liujinzhao@rails.cn

Wenxuan Zhang
Infrastructure Inspection Research Institute
China Academy of Railway Sciences Corporation Limited
Beijing, China
zhangwenxuan@rails.cn

Xiaodi Xu
Infrastructure Inspection Research Institute
China Academy of Railway Sciences Corporation Limited
Beijing, China
xuxiaodi@rails.cn

Zhipeng Yang
Infrastructure Inspection Research Institute
China Academy of Railway Sciences Corporation Limited
Beijing, China
yangzhp@rails.cn

Abstract—In this article, a new pantograph vibration interference signal identification method based on SVM classification is proposed to deal with the abnormal signals caused by transient electromagnetic interference due to pantograph acceleration sensor's suffering from train over-voltage phase break. Firstly, the deviation position is determined by calculating the catenary impact index (CII) of pantograph vertical acceleration, then the acceleration attenuation coefficient of deviation at the said position is calculated, and finally the CII and acceleration attenuation coefficient is selected to form an identification feature vector, thus establishing an interference signal identification method based on support vector machine (SVM). The results show that such an identification structure has fast computing and high accuracy in judging whether the deviation is transient electromagnetic interference. Applying method to analyze the measured data of catenary inspection vehicles demonstrates its effectiveness.

Keywords—Railway Engineering; Catenary; Support Vector Machine (SVM); Articulated Phase Insulator; Transient Process; Electromagnetic Interference（EMI）

I. INTRODUCTION

Since the Guangzhou-Shenzhen Railway, China has started to promote anchoring section articulated phase insulator on raising-speed trunk railway and high-speed railways to satisfy the smooth passage of pantograph when trains are running at high speed [1]. However, this neutral-section passing mode also exposes new electrical problems.
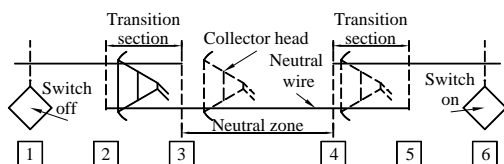
Fig. 1. Schematic diagram of articulated phase insulator

Fig. 1 shows an articulated phase insulator. In the zone between position 2 and position 3, the contact line is suspended parallel to the neutral wire, the pantograph contacts 2 wires at the same time and then transits into the state of contact with the neutral wire only after a period of time, this zone is therefore called a transition section. Likewise, the zone between position 4 and position 5 is also a transition section. Since there is no electrical connection in the zone between position 3 and position 4, the zone is a neutral zone. The collector head is in the locomotive idle running state when passing through positions 2, 3, 4 and 5. After collector head passes through positions 1 and 6, a transient process will occur due to the off-grid power outage and grid-reconnected live switching of the electric locomotive or EMUs, thus forming electromagnetic interference [2], which will greatly affect the detection results of catenary inspection vehicle. When these interference signals exceed the set threshold and are misjudged as deviations, they will cause waste of manpower and material resources while lowering the efficiency of catenary maintenance. Therefore, the detection and identification of interference signals in pantograph vibration signals in the field is extremely important to eliminate abnormal data and improve the validity of data.

To solve this problem, Y. Gong made overvoltage mechanism analysis [3], performed system modeling for each working state in the process of idle running passing phase break after power failure of electric locomotives, and carried out overvoltage simulation for the transient process occurring during the transition of different working states. Based on analysis of the transient process for idle running passing phase break, the literature [4,5] adds the analysis of transient process occurring in case of the off-grid power outage and grid-reconnected live switching of electric locomotives. These literatures all analyze the transient process of neutral-section passing from the perspective of electricity. However, the transient process will also cause serious electromagnetic interference to pantograph acceleration sensors close to the contact line; so the impact acceleration of collector head becomes the main parameter for dynamic catenary detections [6]. This will generate numerous electromagnetic interference signals in pantograph acceleration detection data.

For interference signal processing, relevant researchers have done a lot of researches. For example, L. Yuan et al. made the signal-noise separation against data and extracted the interference signals in the spectrum by using binary

---

* Jinzhao Liu is the corresponding author. (e-mail: liujinzhao@rails.cn).

methods; but the signal-noise separation method selected thereof has relatively poor detection capability for weak signals [7]. G. Antonini proposed a wavelet packet-based interference signal feature extraction method [8]. P.J. Moore [9] and A. Mariscotti [10] respectively analyzed the frequency-domain characteristics and time-frequency characteristic of electromagnetic interference, but ignored how to identify electromagnetic interference in time domain. M.A. Azpurua et al. separated the main components of complex electromagnetic interference by combining empirical mode decomposition (EMD) and transient mode decomposition [11], but this method still fails to satisfy the on-line application of engineering given that the EMD requires a lot of time.

To identify the running state of equipment, L. Wu extracted multiple parameters of vibration signals as SVM feature vectors and classified several samples. However, too many feature vectors brings forth high computational complexity during processing a large amount of data, which does not satisfy the on-line application of engineering [12].

Based on the above analysis, this article proposes to determine the deviation position by calculating the CII followed by the calculation of acceleration attenuation coefficient of each deviation position, and train the model based on SVM classification, with samples that are known to be electromagnetic interference or not as training sets, and the maximum acceleration amplitude and the acceleration attenuation coefficient as prediction variables, to judge whether each sample is electromagnetic interference.

II. IMPACT INDEX METHOD

Given that the pantograph vertical acceleration is profoundly random, the CII method is adopted to demodulate the catenary short wave impact from high frequency to high-stability low-frequency signal, which is more conducive to determining the location of catenary short wave defects.

Detailed calculation steps of catenary short wave impact index method are as follows:

1)   Calculate the effective displacement value of pantograph vertical vibration acceleration:

$$S_r = \sqrt{\sum_{i=r}^{r+K-1} x_i^2 / K} \qquad (1)$$

Wherein: K is the windowing length of the forward windowing. $x_i (i = 1, 2, \ldots, N)$ is the set of filtered waveform signals of pantograph vertical vibration acceleration, and N is the number of waveform signals of pantograph vertical vibration acceleration;

2)   Divide the catenary into several units, with unit length generally taken as 50m;

3)   Calculate the maximum value $S^{\max}$ of effective displacement values of each unit and record it as the unit effective value;

4)   Calculate the average value of the unit effective values for all railway lines of the same speed grade and record it as the calibration parameter $\overline{S}$;

5)   Calculate the CII

$$CII = \frac{S_r}{\overline{S}} \qquad (2)$$

6)   Judge whether the threshold is surpassed, and record corresponding position information [13].

The CII is adopted to analyze the pantograph vertical acceleration. The original signal of pantograph vertical acceleration and the CII of contact line are shown in 0 Comparison between 0(a) and 0(b) shows that the amplitude of the pantograph vertical vibration acceleration signal is quite random, making it difficult to determine the evaluation threshold. However, an independent large peak value can be observed in the CII diagram of pantograph vibration characteristics depicted in terms of energy in Fig. 2(c) and Fig. 2(d), the waveforms are similar, and the distribution characteristics of evaluation indexes after normalization are the same. This indicates that the evaluation method proposed in this article is stable and has good repeatability; by using its distribution law, it is naturally easier to determine the evaluation threshold, which is taken as 4.0 upon calculation and analysis of numerous test data.
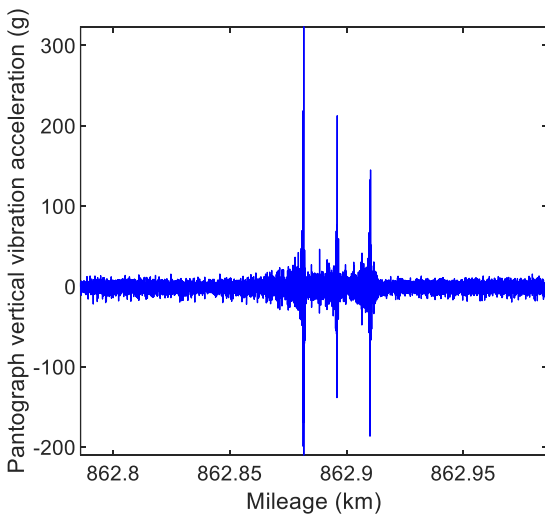
（a）Original Waveform of First Passing
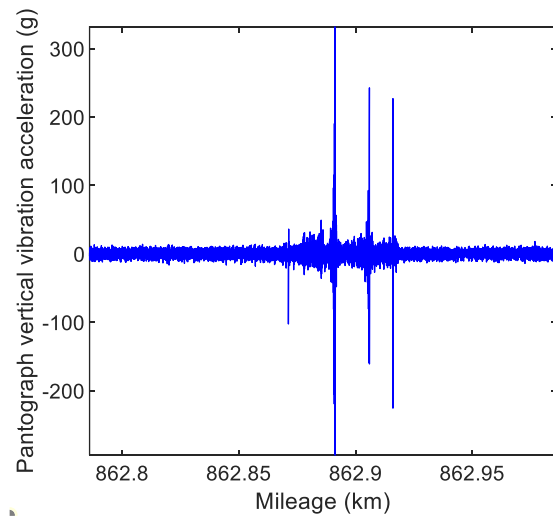
（b）Original Waveform of Second Passing
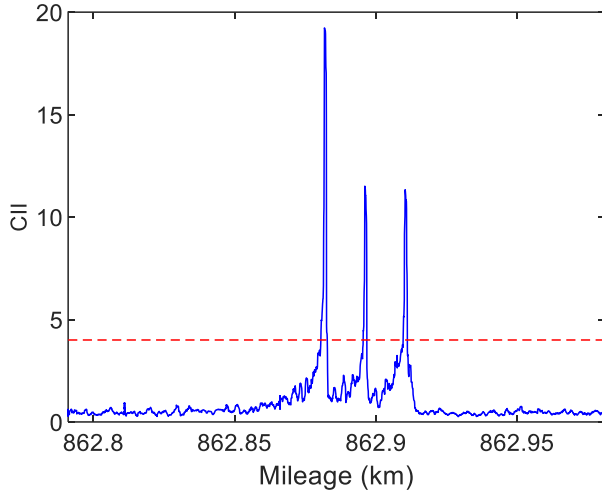
(c) CII of First Passing (d) CII of Second Passing
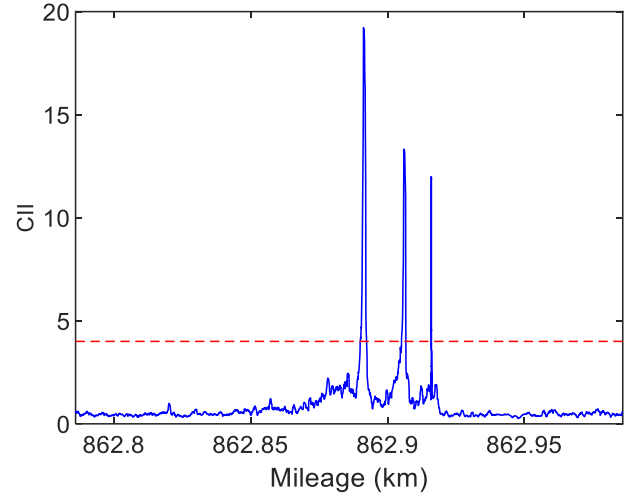
Fig. 2. The pantograph vertical acceleration and its CII of pantograph passing through K862+886 twice.

III. ACCELERATION ATTENUATION COEFFICIENT

Although the CII has good repeatability, it is difficult to judge whether a large peak value is caused by objective defects or by interference signals. As shown in Fig. 3 and Fig. 4, the CIIs at defects and interference signals are far beyond the threshold, and the data characteristics are similar. Therefore, it is difficult to judge whether the deviation is caused by interference signal based on the CIIs alone, and the original signal data of the two cases need to be analyzed separately.
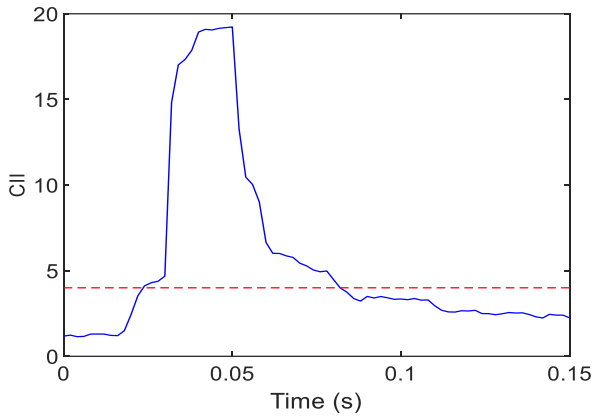
Fig. 3. CII at Defects
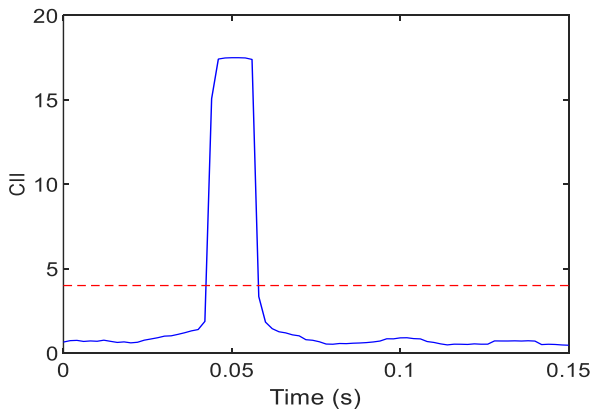
Fig. 4. CII at Interference Signals

When the pantograph passes through the objectively existing defects, the amplitude of the original signal will gradually decrease as shown in Fig. 5, and the vibration will decay to a normal small range in a period of time only after overcoming the external resistance acting.
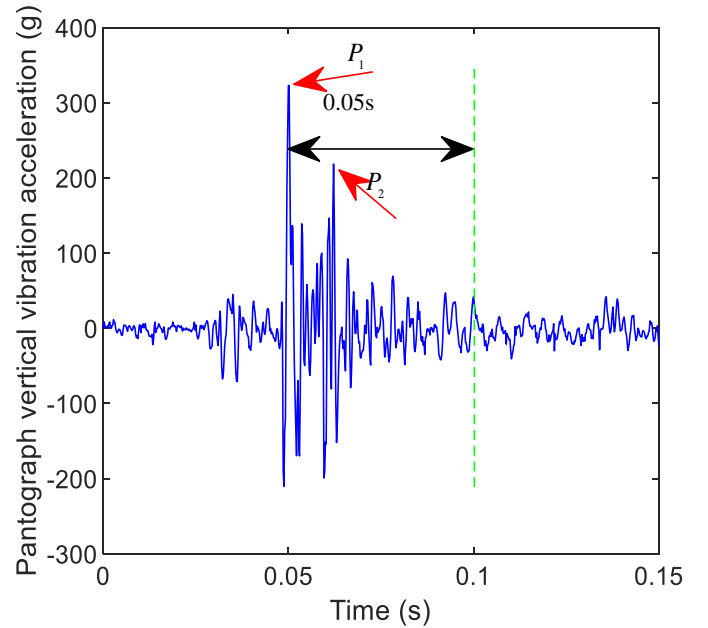
Fig. 5. Pantograph vertical vibration acceleration attenuation waveform of diseases

When the deviation waveform is caused by electromagnetic interference, the instantaneous electromagnetic interference may cause an instantaneous abnormal maximum value of the sensor as shown in Fig. 6, but the value will instantly return to normal level. This abnormality characteristic is different from the data characteristic of objectively existing catenary short wave irregularity.

114
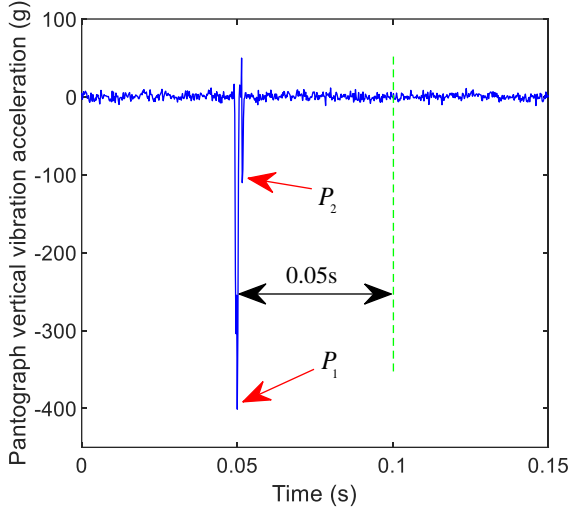
Fig. 6. Pantograph vertical vibration acceleration attenuation waveform of interference signal

Based on the different data characteristics as shown in Fig. 5 and Fig. 6, the ratio of the maximum amplitude P_1 of pantograph vertical vibration acceleration at the deviation position to the maximum amplitude P_2 within a certain time after the acceleration data passes through x-coordinate (based on the calculation and analysis of numerous test data, the time horizon is taken as 0.05s) is calculated and recorded as the acceleration attenuation coefficient R:

$$R = \frac{P_1}{P_2} \tag{3}$$

IV. PANTOGRAPH VIBRATION INTERFERENCE SIGNAL IDENTIFICATION METHOD BASED ON SVM CLASSIFICATION

A. Principle of SVM Classification

Support Vector Machine (SVM) is a kind of generalized linear classifier, which classifies data by supervised learning, with the maximum-margin hyperplane for learning sample solution as decision boundary [14-16].

Input data and learning objectives $X = \{X_1, \ldots, X_N\}, y = \{y_1, \ldots, y_N\} \left(X_i \in R^d; y_i \in \{-1, +1\} \right)$ are given. If the samples are linearly separable, the support vector machine will transform the classification problem into the solution of convex quadratic optimization:

$$\begin{cases} \min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} \xi_i \\ s.t.\ y_i \left(\omega^\top \cdot X_i + b \right) \geqslant 1 - \xi_i \end{cases} \tag{4}$$

Wherein: ω is a weight vector;

C is the penalty factor;

ξ is the relaxation factor;

b is the offset.

The optimal classification decision function obtained is as follows:

$$sgn\left[y_i \left(\omega^\top \cdot X_i + b \right) \right] \tag{5}$$

When the order of polynomial kernel is not 1, nonlinear SVM can be obtained.

For nonlinear SVM optimization problem:

$$\begin{cases} \min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} \xi_i \\ s.t.\ y_i \left[\omega^\top \cdot \phi(X_i) + b \right] \geqslant 1 - \xi_i, \xi_i \geq 0 \end{cases} \tag{6}$$

The optimal classification decision function obtained is as follows:

$$sgn\left[y_i \left(\omega^\top \cdot \phi(X_i) + b \right) \right] \tag{7}$$

B. SVM-Based Interference Signal Recognition

According to the calculation and analysis of numerous test data, the CII of interference signal often far exceeds the threshold, but there are also a few serious defects where the CII far exceeds the threshold, and even produce a peak value of the same size as the impact index of interference signal. Therefore, it is difficult to determine whether the deviation is caused by interference signal based on CII alone. For defects of varying severity, the distribution of the acceleration attenuation coefficient of the original signal is also different. Therefore, CII and acceleration attenuation coefficient proposed in Sections 1 and 2 are assumed to be signal feature vectors as input samples for SVM classification. The specific process may be found in Fig. 7.
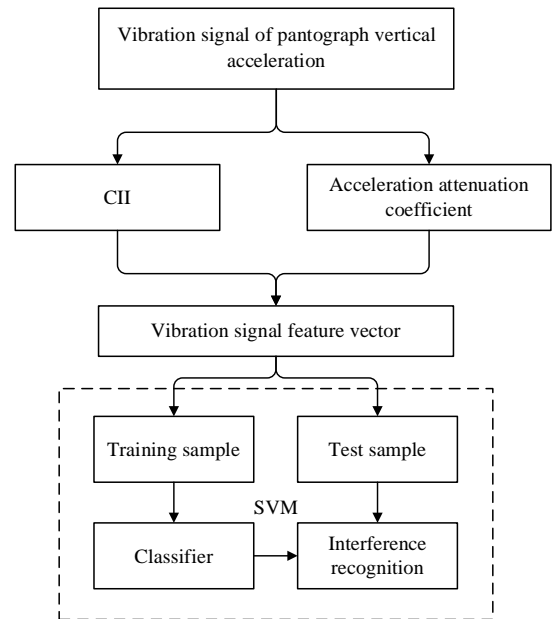
Fig. 7. Flow chart of interference signal identificatio

115

As shown in Fig. 7, acceleration attenuation coefficient are extracted firstly based on the vibration signal of pantograph vertical acceleration, and the CII and the acceleration attenuation coefficient are taken as vibration signal feature vector. Multiple groups of vibration signal feature vectors extracted to determine whether they are interference signals are respectively taken as SVM classification training samples and test samples. Training samples are trained by principles of empirical risk and structural risk minimization, to obtain decision functions between feature vectors and interference state parameters. Then the same principles are further applied to the test samples to obtain the accuracy rate of test sample classification results and model prediction.

V. CASE ANALYSIS

The proposed method is used to analyze the date of pantograph vertical acceleration for multiple railway lines.

CII of pantograph vertical acceleration on each railway line is calculated, 142 vibration signals exceeding the management value of catenary impact index are selected as samples, to judge through manual review whether each deviation is a true defect; if so, the state function value of this group of samples is 1. If no defect is found in the field review but an instantaneous maximum value appears in the waveform, it means that this place accommodates an interference signal, and then the state function value of this group of samples is-1. After the state function values of each group is determined, all samples are divided into training samples and test samples, including 111 groups of training samples and 31 groups of test samples.

CII and acceleration attenuation coefficient of training samples are set as prediction variables, the state function used to judge whether the sample is an interference signal is set as a response, and the 10-fold cross validation is adopted to estimate the algorithm accuracy.
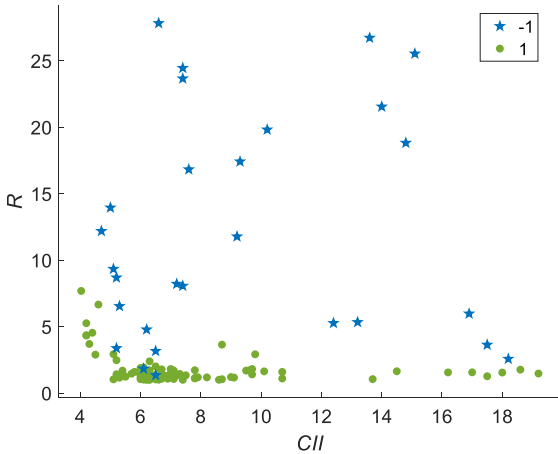
Fig. 8. Original data set

Different categories in the Original data set are distributed as shown in Fig. 8; from which, the distribution area of Class-1 data points is obviously different from that of Class 1 data points.

Various SVM classification methods are used to train the training data, obtaining the accuracy of each model as shown in TABLE I. :

TABLE I.   ACCURACY OF VARIOUS SVM MODELS

Different SVM	Linear	Quadratic function	Cubic function
Accuracy	93%	92%	93%
Different SVM	Fine Gaussian	Medium Gaussian	Coarse Gaussian
Accuracy	95%	90%	87%

From Table 1, the model trained with the fine Gaussian SVM classification method has the highest accuracy, followed by the cubic SVM and Linear SVM.

Therefore, fine Gaussian SVM classification method is chosen to train the training set.
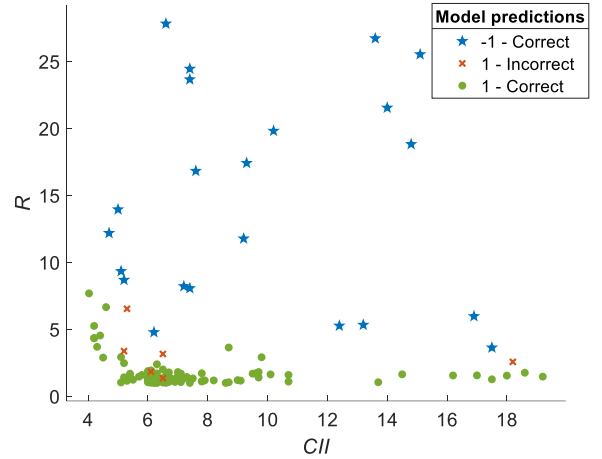
Fig. 9. Model trained with fine Gaussian SVM

Crosses in Fig. 9 represent incorrect predictions, dots and five-pointed stars represent correct predictions, while there are only a few crosses in the whole scatter diagram, which represents that the trained model has high prediction accuracy. Such a result is consistent with that shown in Table 1.
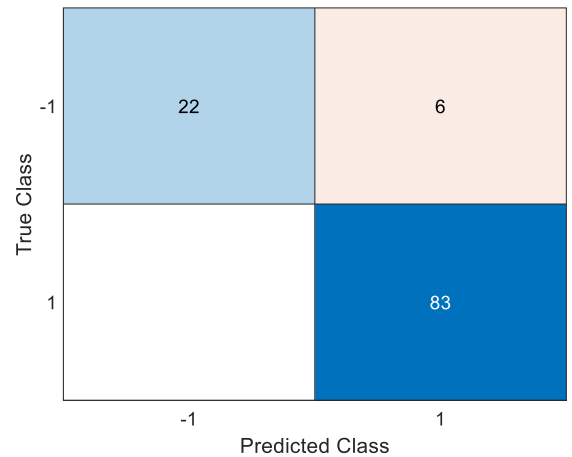
Fig. 10. Confusion matrix of the model

Four regions in Fig. 10 are in turns TN: true negative sample, FP: false positive sample, FN: false negative sample and TP: true sample. Green areas in the upper left corner and the lower right corner represent the correct prediction of samples, while the areas in the lower left corner and the upper

right corner represent the incorrect prediction. It can be observed from the figure that only 6 negative samples are misjudged as positive samples. Among them, 4 samples are due to two occurrences of electromagnetic interference with the same amplitude in the range of 0.05s, 1 sample is due to the small electromagnetic interference amplitude, and the sample with the largest CII is due to electromagnetic interference that happened to be near a disease with a large vibration amplitude. These conditions with incorrect predictions are relatively rare.

TPR and FPR can be calculated with values of the four parameters TP, FN, FP and TN in the confusion matrix shown in Fig. 10:

$$
\begin{cases}
TPR = \dfrac{TP}{TP + FN} \\
FPR = \dfrac{FP}{FP + TN}
\end{cases}
\tag{8}
$$
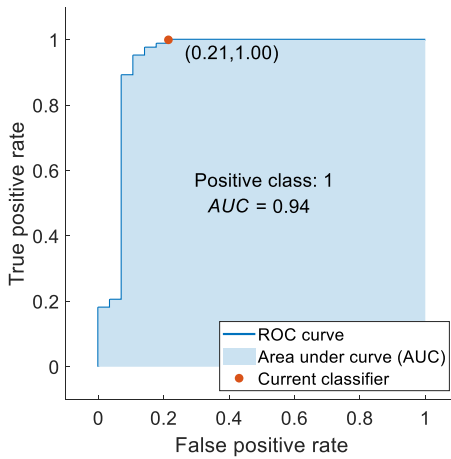
Fig. 11. ROC curve of the model

The ROC curve drawn with TPR and FPR is shown in Fig. 11. The area enclosed by the ROC curve is AUC; and AUC of the model shown in the figure is 0.94, indicating that the model has a very good classification effect.

The model is exported, and prediction variables CII, R and responses of test samples are imported into the model to obtain predicted responses, which is calculated and verified with the responses of known categories, and the accuracy obtained is 96.77%.

According to the scatter diagram, confusion matrix, AUC calculated with ROC curve, and the accuracy of test samples, the model has a higher accuracy in judging whether the deviation is transient electromagnetic interference.

VI. CONCLUSIONS

To identify interference signals in the pantograph vibration signals, this article proposes a method of identifying interference signals based on pantograph vibration signal SVM classification. Under the method, the deviation position is determined by calculating the CII, and then the acceleration attenuation coefficient of each deviation position is calculated, while the model is trained based on SVM classification, with samples that are known to be electromagnetic interference or not as training sets, and the CII and the acceleration attenuation coefficient as prediction variables. Through example analysis, evaluation of multiple indexes and verification of test sets, the results show that the model has a high accuracy, and that the method can well judge whether the deviation is transient electromagnetic interference, better guide the maintenance of catenary, and satisfy the on-line application of engineering.

REFERENCES

[1] W. Yu, Catenary System of High-speed Electrified Railway, Southwest Jiaotong University Press, 2003.

[2] Q. Li, and J. He, Traction power system analysis, Southwest Jiaotong University Press, 2012.

[3] Y. Gong, Research of Over-voltages of Electric Locomotive Passing the Articulated Phase Insulator, Journal of the China Railway Society, no. 04, 2008, pp. 103-107.

[4] N. Li, Research on Electromagnetic Transient Process of Electrical Locomotive System, Beijing Jiaotong University, 2010.

[5] X. Jiang, Z. He, H. Hu, S. Gao, and B. Wang, Analysis on Electromagnetic Transient Process of Electric Multiple Unit Passing Neutral Section Devices, Journal of the China Railway Society, vol. 35, no. 12, 2013, pp. 30-36.

[6] X. Kang, J. Liu, T. Han, and Z. Sun, Research on the Criteria for Evaluating the Hard Spots of Catenary Based on the Vertical Impact Acceleration of Collector head, China Railway Science, vol. 30, no. 01, 2009, pp. 108-113.

[7] L. Yuan, Q. Liu, Z. Sun, Y. Liu, Y. Wang, F. Liu, and M. Chen, Detection and identification for electromagnetic interference of equipment area on-site, Chinese Journal of Radio Science, vol. 32, no. 06, 2017, pp. 650-656.

[8] G. Antonini, and A. Orlandi, Wavelet packet-based EMI signal processing and source identification, Ieee Transactions On Electromagnetic Compatibility, vol. 43, no. 2, 2001, pp. 140-148.

[9] P.J. Moore, and V.S.H. Chong, Analysis of the radiated electromagnetic field generated by a 132 kV, SF6 circuit breaker, Conference on Electrical Insulation & Dielectric Phenomena. IEEE, 2002.

[10] A. Mariscotti, A. Marrese, N. Pasquino, and R. Schiano Lo Moriello, Time and frequency characterization of radiated disturbance in telecommunication bands due to pantograph arcing, Measurement, vol. 46, no. 10, 2013, pp. 4342-4352.

[11] M.A. Azpurua, M. Pous, and F. Silva, Decomposition of Electromagnetic Interferences in the Time-Domain, Ieee Transactions On Electromagnetic Compatibility, vol. 58, no. 2, 2016, pp. 385-392.

[12] L. Wu, Q. Wu, Y. Feng, Z. Yao, Z. Wang, and L. Zhao, State Recognition of Transformer Based on SVM Classification of Vibration Signals, High Voltage Apparatus, vol. 55, no. 11, 2019, pp. 232-238.

[13] Y. Ding, X. Xu, J. Liu, and W. Zhang, Impact Index Method for Railway Catenary Short-Wave Irregularity Evaluation, Structural Health Monitoring 2019, Year Published.

[14] V. V, Statistical learning theory Wiley, John Wiley & Sons,Inc, 1998.

[15] Z. Zhou, Machine Learning, Tsinghua University Press, 2016:pp.121-139, 298-300.

[16] H. Li, Statistical Learning Method, Tsinghua University Press, 2012:pp.95-135.

Coupling Deep Models and Extreme Value Theory for Open Set Fault Diagnosis

Xiaolei Yu
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an,China
yxl007@stu.xjtu.edu.cn

Zhibin Zhao
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an,China
zhibinzhao1993@gmail.com

Xingwu Zhang*
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an,China
xwzhang@mail.xjtu.edu.cn

Chuang Sun
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an,China
ch.sun@xjtu.edu.cn

Qiyang Zhang
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an,China
zhangqiyang@stu.xjtu.edu.cn

Xuefeng Chen
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
chenxf@mail.xjtu.edu.cn

Abstract—**Existing deep-learning-based fault diagnosis methods assume that all possible fault modes are available during training process, which is sometimes not consistent with real applications. Unknown fault types may occur in the testing phase due to the fact that it is impossible to collect all the fault modes in the training phase. Thus, in this paper, we introduce and define the open set fault diagnosis (OSFD), and handle this problem in both shared-domain and cross-domian scenarios. For shared-domain OSFD, an extreme-value-theory-based method is proposed to build a rejection model to detect samples from the unknown classes. For cross-domain OSFD, weighted domain adversarial neural networks is constructed to obtain domain-invariant features of the shared classes and separate samples of unknown classes by reweighting target samples. Learned features of source data are used to establish a rejection model, such that unknown samples in the target domain can be detected. Experimental results on the Case Western Reserve University dataset demonstrate the effectiveness of the proposed methods.**

Keywords—Domain adaptation, open set fault diagnosis, extreme value theory, weighted domain adversarial neural networks.

I. Introduction

Recently, deep-learning-based methods have been widely used to tackle the machinery fault diagnosis problem due to their powerful representation learning ability, and two open-source benchmark studies were performed in [1, 2]. These methods perform well under the hypothesis that labeled data with variable fault modes are abundant and the machine works under a constant operation condition.

However, the location, type, extent, and number of occurred faults are uncertain in the real situation and it is not realistic to collect data with all kinds of fault types. For example, bearing faults may occur on the inner race, outer race, roller, and cage, but the fault symptom (pitting, indentations and flaw) and the fault distribution (local fault and distributed fault) may be different. Hence, it is inevitable that some fault types of test data are not included in training data, which can be regarded as an open set fault diagnosis (OSFD) problem. These fault types of test data unseen in training data are denoted as unknown class. Corresponding test data are denoted as unknown samples. In view of this problem, two open set fault diagnosis settings are considered in this paper:

1) Shared-domain open set fault diagnosis (SOSFD): Training data and test data are collected under the same operation condition. The fault modes of training data are a subset of test data.

2) Cross-domain open set fault diagnosis (COSFD): training data (source data) and test data (target data) are collected under different operation conditions. The fault modes of source data are a subset of target data.

OSFD is an open set recognition (OSR) [3] problem. There have been already a variety of works for OSR [4]. Bendale et al. [5] defined the values in the penultimate layer of neural network as the activation vector (AV) and proposed to replace the SoftMax layer with an OpenMax layer. AV is analyzed based on extreme value theory (EVT) to evaluate if test samples are far from training samples. This work opened up a novel idea for the subsequent research [6-8]. Thus far, There is limited research covering OSFD. Tian et al. [9] proposed to tackle the OSFD problem by combining k-nearest-neighbor-search-based local learning with kernel null Foley-Sammon transform. Chao et al. [10] considered a semi-supervised scenario where only healthy data are labeled and data with different fault types need to be distinguished. By combining modified variational autoencoders with an one-class classifier, the latent representation space of healthy data is restricted. Then the latent representations of unlabeled fault data are clustered by a density-based clustering algorithm. However, these works mainly focus on SOSFD, which neglect the effect of the domain shift.

In the real industrial application, the operation conditions of machines vary frequently, which means the distributions of training data and test data are different and results in a domain shift phenomenon. Open set domain adaptation (OSDA) [11] methods aim to tackle the OSR problem when the training data and test data come from different distributions. Limited studies can be found in the current literature on OSDA. Saito et al. [12] extended the OSR problem to the setting of deep OSDA for the first time. A feature generator and classifier are trained adversarially to make a decision boundary for unknown class. Liu et al. [13] added a multi-binary classifier and a binary classifier to the adversarial domain adaptation network to produce weights to rejecte target samples from the unknown classes, thereby samples of the unknown classes can be separated from other samples. To our best knowledge, there is no study on the COSFD which has practical significance.

For these two settings, the common goal is that samples of known classes are correctly classified while samples of unknown class are rejected. For the SOSFD, we follow the EVT based approach [5]. EVT is applied to analyze the feature

distribution of each class in the training set and a rejection model is established to decide if a test sample should be rejected. Test samples that are not rejected are classified to known classes according to their SoftMax probabilities. For the COSFD, EVT-based OSR methods may fail to work. Since the feature distribution of target data is far away from source data, which causes the rejection of most of target data. Consequently, weighted domain adversarial neural networks is proposed to obtain domain-invariant features on the shared class of source and target domains, and to separate unknown samples of target domain. After that, the aforementioned EVT-based OSR method is applied to the learned features of target data.

The remainder of this paper is summarized as follows: In Section II, the preliminary is illustrated. Section III presents the details of the proposed OSFD approach. Detailed experiments and analyses are carried out in Section IV. At last, the conclusions are drawn in Section V.

II. PRELIMINARY

A. EVT based Open Set Recognition

Scheirer et al. [14] analyzed the distributions of the recognition system extreme scores based on EVT and found these distributions follow a Weibull distribution. When applied to latent representations in deep models, EVT based open set recognition has been shown to effectively detect out-of-distribution (OOD) samples [15]. After training a feature extractor and a classifier, the latent representations and label predictions can be obtained. The latent representations for each training sample and test sample are denoted as $\{f_i^L\}_{i=1}^{n_L}$ and $\{f_i^U\}_{i=1}^{n_U}$, respectivly, where n_L is the number of training samples and n_U is the number of test samples. For each class c, let $S_{i,c}^L = f_{i,c}^L$ denote the latent representations of correctly classified training samples. The per class mean latent representation u_c^L is computed by averaging $S_{i,c}^L$. Next, distances between all correctly classified training samples and the associated u_c^L in the same class are calculated as follows:

$$d_c = \left\| S_c^L - u_c^L \right\|, \quad (1)$$

By fitting a Weibull distribution using d_c and tail-size μ, the Weibull distribution ρ_c whose parameters include shift τ_c, shape κ_c and scale λ_c are returned for each class. At test time, the class of each test sample is computed according to its label prediction. Then distances between all test samples and the associated u_c^L in the same class are calculated as follows:

$$d_c^{'} = \left\| u_c^L - f_{i,c}^U \right\|, \quad (2)$$

$d_c^{'}$ is used to calculate the Weibull cumulative distribution function (CDF) probability:

$$\omega_c\left(d_c^{'}\right) = 1 - \exp\left(-\frac{\left\| d_c^{'} - \tau_c \right\|}{\lambda_c}\right)^{\kappa_c}, \quad (3)$$

The Weibull CDF probability represents the probability that a test sample come from the unknown class. If this probability exceeds a predefined threshold, this sample is considered as an unknown sample.

The corresponding procedures to fit Weibull models and compute Weibull CDF probability are summarized in Algorithms 1 and Algorithms 2, respectively.

Algorithm 1: Establish EVT Model Through Deep Neural Network for Spen set Recognition

---

Input: latent representations of labeled training set $\{f_i^L\}_{i=1}^{n_L}$, tail-size μ

For each class c, let $S_{i,c}^L = f_{i,c}^L$ for each correctly classified training example

for $c = 1 \ldots n_c$ do

    Compute per class mean latent representations, $u_c^L = mean(S_{i,c}^L)$

    Weibull model $\rho_c = (\tau_c, \kappa_c, \lambda_c)$ =Fit Weibull $\left(\left\| S_c^L - u_c^L \right\|, \mu\right)$

Algorithm 2: Open Set Probability Estimation for Unknown Inputs

---

Input: latent representations of unlabeled test set $\{f_i^U\}_{i=1}^{n_U}$, Per class latent mean u_c^L and Weibull model ρ_c, each with parameters $(\tau_c, \kappa_c, \lambda_c)$, threshold Ω

Compute distances to u_c^L : $d_c^{'} = \left\| u_c^L - f_{i,c}^U \right\|$

for $c = 1 \ldots n_c$ do

    Weibull CDF $\omega_c\left(d_c^{'}\right) = 1 - \exp\left(-\frac{\left\| d_c^{'} - \tau_c \right\|}{\lambda_c}\right)^{\kappa_c}$

Reject input if $\omega_c\left(d_c^{'}\right) > \Omega$ for any class c

---

B. Domain Adversarial Neural Networks

Domain adversarial neural networks (DANN) [16] is a state-of-the-art domain adaptation method which has achieved excellent performance on various image datasets. Therefore, this method is used as the basis of the proposed COSFD method. Pioneered by the idea of Generative Adversarial Networks (GANs) [17], DANN exploited a domain classifier denoted by D to obtain domain invariant features. D is a binary domain classifier with all the source samples labeled as 1 and all the target samples labeled as 0. The domain classifier D is trained to distinguish source samples from target samples while the feature extractor F is trained to generate features that can fool D. Specifically, the parameters of the feature extractor F are trained to maximize the loss of the domain classifier D and minimize the loss of the label predictor G. But the parameters of domain classifier D are trained to minimize its loss. In this way, the distribution discrepancy between source and target domains can be alleviated.

Let $f = F(x)$ denotes feature learned by features extractor.

The loss of the domain classifier L_d is calculated as follows:

$$L_d = -\mathbb{E}_{x_i^s \sim D_s} \log\left[D\left(f_i^s\right)\right] - \mathbb{E}_{x_j^t \sim D_t} \log\left[1 - D\left(f_j^t\right)\right]. \quad (4)$$

where $\mathbb{E}$ denotes the expectation, f_i^s denotes the features of the i th source sample, and $D\left(f_i^s\right)$ is a binary variable (0 or 1) to indicate the domain label of features.

Moreover, the parameters of the feature extractor F and the label predictor G are trained to minimize the loss L_g using source domain data to obtain discriminative features:

$$L_g = -\mathbb{E}_{(x_i^s, y_i^s) \sim D_s} \sum_{c=1}^{n_c} 1_{[y_i^s = c]} \log \left[G_c \left(f_i^s \right) \right], \quad (5)$$

where n_c denotes the number of categories in source data.

The overall objective functions are as follows:

$$\max_{F,G} -L_g + \lambda_1 L_d$$
$$\min_D L_d, \quad (6)$$

where λ_1 is a tradeoff parameter.

III. OPEN SET FAULT DIAGNOSIS

In this section, the problem description and the proposed OSFD approach are illustrated in detail. The overall diagnosis procedure is shown in Fig. 1.
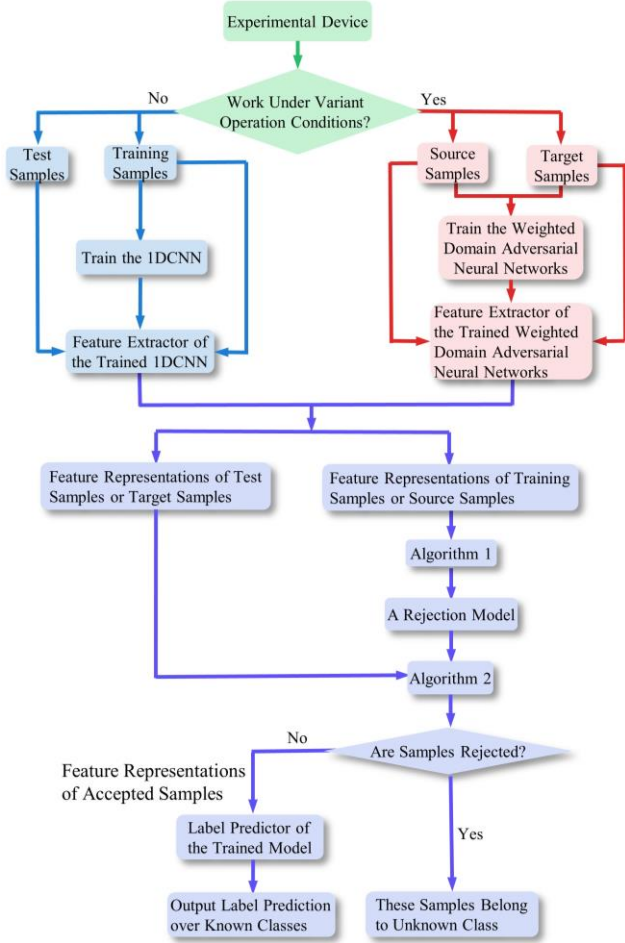
Fig. 1. The diagnosis procedure of open set fault diagnosis.

A. Problem Description

Considering that fault diagnosis approaches need to identify the data of entirely new (unknown) fault modes in practice, two problems are studied in this paper:

1) SOSFD: Training data and test data are collected under the same operation condition. The fault modes of training data are a subset of test data. Training data denoted by $D_L = \{(x_1^L, y_1^L), \cdots, (x_{n_L}^L, y_{n_L}^L)\}$ consists of n_L samples of K classes. Test data denoted by $D_U = \{x_1^U, \cdots, x_{n_U}^U\}$ consist of n_U samples of $M > K$ classes, including the known classes in addition to one or more unknown classes.

2) COSFD: Source data and target data are collected under different operation conditions. The fault modes of source data are a subset of target data. Source data denoted by $D_s = \{(x_1^s, y_1^s), \cdots, (x_{n_s}^s, y_{n_s}^s)\}$ consists of n_s samples of K classes. Target data denoted by $D_t = \{x_1^t, \cdots, x_{n_t}^t\}$ consist of n_t samples of $M > K$ classes, including the known classes in addition to one or more unknown class.

The goal of the proposed methods is to correctly classify samples of $K+1$ classes where samples of K classes are classified into corresponding classes and samples of the $M-K$ unknown class are assigned to the additional classes.

B. Shared-domain Open Set Fault Diagnosis Approach

For the SOSFD, the network architecture is a one-dimensional convolutional neural network (1DCNN) which is composed of a feature extractor F and a label predictor G. The label predictor is a K-class classifier.

In previous studies, 1DCNN performed well in closed set fault diagnosis. The feature extractor and the label predictor are trained to obtain discriminative features and recognize fault modes. However, for the OSFD, test data contain some unknown class that do not appear in training data. 1DCNN is not capable to detect these unknown samples because the Softmax function only generates a probability distribution over the K known class labels. To tackle this problem, features of training data are input to Algorithm 1 to build a rejection model. During testing, the Weibull CDF probability of each test sample is computed (according to Algorithm 2) to give the rejection probability. A samples is considered as unknown sample if its Weibull CDF probability exceeds a threshold Ω. Samples that are not rejected are input to the label predictor to produce probabilities over known classes.

C. Cross-domain Open Set Fault Diagnosis Approach

For the COSFD, besides the aforementioned feature extractor F and the label predictor G, two domain classifiers D and D_ω which are composed of fully connected layers are added to the network architecture as shown in Fig. 2.
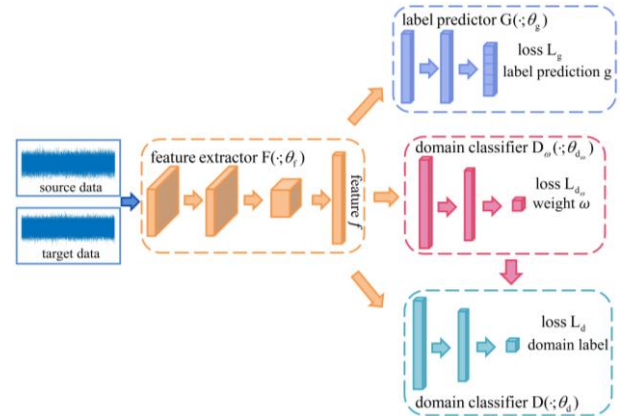
Fig. 2. Illustration of the network architecture of weighted domain adversarial neural networks.

Due to that source data and target data are collected under different operation conditions. The aforementioned algorithm will not work due to the huge distibution discrepancy between source and target domians. A natural idea to solve this problem is to use existing domain adaptation methods such as DANN. But these methods align the whole target domain with source domain without considering samples from unknown classes, which may lead to class mismatch problem.

Inspired by the idea of [18], weighted domain adversarial neural networks (WDANN) is proposed to weight the importance of target samples on feature distribution alignment which can match samples of shared classes better and separate samples of unknown and known classes. Another domain classifier D_ω, whose output value gives the likelihood of a sample coming from source domain, is added to the architecture of DANN. Let $f = F(x)$ denotes feature learned by features extractor. If $D_\omega(f) \approx 0$, the sample is highly likely to come from the unknown classes in the target domain, since the sample can be perfectly distinguished from the source domain. The contribution of these samples should be small such that both the domain classifier and the feature extractor will ignore them. On the other hand, if $D_\omega(f) \approx 1$, the sample is more likely from the shared classes of target domain. These samples should be given a larger weight to reduce the distribution discrepancy between the shared classes. Hence, $D_\omega(f)$ can be used to define the weight of target sample in the domain classifier D. The weight is defined as follow:

$$\omega(f) = D_\omega(f). \tag{7}$$

By applying the weights on the target features in the domain classifier D, the loss L_d is modified to:

$$L_d = -\mathbb{E}_{x_i^s \sim D_s} \log\left[D\left(f_i^s\right)\right] - \mathbb{E}_{x_j^t \sim D_t} \omega\left(f_j^t\right) \log\left[1 - D\left(f_j^t\right)\right]. \tag{8}$$

In this way, the weights of target samples from unknown class will be smaller than those from the shared classes and the domain discrepancy between the shared classes will be reduced.

The domain classifier D_ω is only used for generating the weights of target samples. Thus the loss L_{d_ω} of D_ω is only used for updating its parameters but not used for updating the parameters of the feature extractor:

$$L_{d_\omega} = -\mathbb{E}_{x_i^s \sim D_s} \log\left[D_\omega\left(f_i^s\right)\right] - \mathbb{E}_{x_j^t \sim D_t} \log\left[1 - D_\omega\left(f_j^t\right)\right]. \tag{9}$$

The overall objective functions are as follows:

$$\max_{F,G} - L_g + \lambda_1 L_d$$
$$\min_{D} L_d \tag{10}$$
$$\min_{D_\omega} L_{d_\omega},$$

where λ_1 is the tradeoff parameter.

It is convenient to train the WDANN by the stochastic gradient descent (SGD) with a momentum algorithm. Let $\theta_f, \theta_g, \theta_d$ and θ_{d_ω} be the parameters of the feature extractor F, the label predictor G, the domain classifier D, and the domain classifier D_ω, respectively. These parameters are updated as follows:

$$\theta_f \leftarrow \theta_f - \eta \frac{\partial\left(L_g - \lambda_1 L_d\right)}{\partial \theta_f} \tag{11}$$

$$\theta_g \leftarrow \theta_g - \eta \frac{\partial L_g}{\partial \theta_g} \tag{12}$$

$$\theta_d \leftarrow \theta_d - \eta \frac{\partial L_d}{\partial \theta_d} \tag{13}$$

$$\theta_{d_\omega} \leftarrow \theta_{d_\omega} - \eta \frac{\partial L_{d_\omega}}{\partial \theta_{d_\omega}}. \tag{14}$$

To solve the minimax optimization problem, a gradient reversal layer (GRL) [16] is added between the feature extractor F and the domain classifier D.

After training, the learned features of source data are input to Algorithm 1 to establish a rejection model and then Algorithm 2 is used to detect weather features of target data belong to unknown class. Like SOPFD, samples that are not rejected are input to the label predictor to produce probabilities over known classes.

IV. CASE STUDY

A. Dataset Description

The dataset used to evaluate the proposed OSFD method is acquired by the bearing data center of the Case Western Reserve University (CWRU) [19]. It was collected by the accelerometers attached to the drive end of motor housing. For fault bearings, single-point faults generated by electro-discharge machining with three kinds of fault size (7 mils, 14 mils, and 21 mils) were introduced to the inner race, the outer race, and the ball of bearings, respectively. Thus, there are nine fault bearings and one healthy bearing are involved in the experiment. Besides, the vibration signals of these bearings were collected at four kinds of motor loads, i.e., 0, 1, 2, and 3hp. Samples with 1024 points are selected through the original vibration signal. The descriptions of the CWRU dataset are shown in Table I.

TABLE I THE INFORMATION OF THE CWRU DATASET

Class label	Fault mode	Fault size(mil)
0	Healthy	0
1	Ball fault	7
2	Ball fault	14
3	Ball fault	21
4	Inner race fault	7
5	Inner race fault	14
6	Inner race fault	21
7	Outer race fault	7
8	Outer race fault	14
9	Outer race fault	21

For the SOSFD, four fault diagnosis tasks are built: T_0, T_1, T_2, and T_3. To train deep network and establish a reject model, the data with labels 0-6 are randomly divided into four parts: training set, validation set, threshold set (generating threshold for rejection model), and test set. The ratios are 60%, 10%, 10% and 20%, respectively. 20% of the data with labels 7-9 are randomly selected and added to the test set. Thereby the test set include 10 classes in total.

For the COSFD, twelve transfer tasks are built: T_{01}, T_{02}, T_{03}, T_{10}, T_{12}, T_{13}, T_{20}, T_{21}, T_{23}, T_{30}, T_{31}, and T_{32}. For instance, T_{01} denotes that source data are collected under 0hp motor load and target data are collected under 1hp motor load. Source data, which are composed of samples with 0-6 labels, are randomly divided into three parts: training set, validation set, and threshold set. The ratios are 80%, 10%, and 10%, respectively. Target data contain all 10 classes and all of them are involved in training and test.

B. Parameter Settings for Open Set Fault Diagnosis

The mini-batch SGD with momentum of 0.9 is used and the batch size of each domain is set to 128. As for the learning rate η, an annealing strategy according to [20] is adopted. The formula is as follows:

$$\eta = 0.01(1+10p)^{-0.75}, \tag{15}$$

where p represents the training progress linearly changing from 0 to 1.

The tail-size is set to 5% of labeled training samples and the threshold is set to the value that ensures 95% of samples under the threshold set not to be rejected as an outlier.

For COSFD, the trade-off parameter λ_1 is gradually changed from 0 to 1 using the following formula:

$$\lambda_1 = \frac{2}{1+\exp(-10p)} - 1. \tag{16}$$

C. Experimental Results

All the experiments are carried out on a PC with Intel Core i5 CPU, 8 GB RAM, and GeForce RTX 2060Ti GPU. the programming platform is Pytorch. Each task is implemented ten times over 200 training epochs and the average experimental results are record to reduce the effect of randomness.

To validate the superiority of the proposed OSFD method, methods without the application of OSR are used for comparison. Baseline_S (in SOSFD) and Baseline_C (in COSFD) are trained and then all test samples (including unknown samples) are classified only according to the label prediction of the model. Following previous works [11, 13], four evaluation metrics are employed for the proposed method, including: ALL: the accuracy for all $K+1$ classes including the unknown as one class; ALL*: the accuracy for all the K known classes; UNK: the accuracy of unknown class; and CS: the accuracy of the K known classes without the application of OSR algorithm. The average classification accuracies for SOSFD and COSFD are listed in Table II and Table III, respectively.

It can be observed that the accuracies of ALL are higher than Baseline_S on all tasks in Table II and the accuracies of ALL are higher than Baseline_C on most of tasks in Table III. This is due to that Baseline_S and Baseline_C are not able to detect unknown samples which means that all unknown samples are false classified, but the proposed method are effective for the rejection of unknown samples according to UNK while remain the capability of known classes classification according to ALL*. We also observe that CS is higher than ALL* on all tasks in both Table II and Table III, which indicates that some samples of known classes are miclassified into unknown class. This phenomenon is more obvious for COSFD since samples from the shared classes in source and target domain cannot match completely. It is inevitable to reject some known samples at the expense of detecting unknown samples but this false rejection should be reduced as possible in future research. More improvement for OSFD is expected.

D. Feature Visualization

T-SNE [21] technology is utilized to visualize the high-

TABLE II CLASSIFICATION ACCURACY (%) OF SHARED-DOMAIN OPEN SET FAULT DIAGNOSIS

Task	Baseline_S	ALL	ALL*	UNK	CS
T_0	72.4	87.4	96.9	62.4	99.5
T_1	76.6	89.4	91.9	81.1	99.7
T_2	76.9	86.9	94.3	62.1	100.0
T_3	76.8	89.8	96.6	67.2	100.0
Average	75.7	88.4	94.9	68.2	99.8

TABLE III CLASSIFICATION ACCURACY (%) OF CROSS-DOMAIN OPEN SET FAULT DIAGNOSIS

Task	Baseline_C	ALL	ALL*	UNK	CS
T_{01}	74.0	82.7	88.4	63.7	95.1
T_{02}	74.6	80.1	87.4	55.4	96.8
T_{03}	69.6	69.1	72.2	59.0	88.9
T_{10}	70.0	81.9	93.0	52.4	97.5
T_{12}	76.6	85.4	96.1	49.6	99.7
T_{13}	74.3	81.0	91.2	46.9	96.5
T_{20}	67.7	67.5	70.4	59.8	93.7
T_{21}	74.4	80.0	88.6	51.8	97.2
T_{23}	74.9	80.2	85.7	61.8	97.6
T_{30}	62.8	60.7	60.2	62.0	85.3
T_{31}	69.4	76.6	80.9	62.7	90.4
T_{32}	76.4	81.9	88.1	61.2	99.5
Average	72.1	77.3	83.5	57.2	94.9

dimensional features. The 2D features of training data and test data on transfer task T_0, T_1, T_2, and T_3 are shown in Fig. 3. It can be observed that most of unknown samples separate from known samples but there are still some unknown samples distribute around known samples. The 2D features of source data and target data on transfer tasks T_{01} and T_{30} are shown in Fig. 4. The feature distributions of shared classes in source and target domains are aligned well on T_{01} but are not aligned sufficiently on T_{30}. It can be inferred that the distribution discrepancy between the shared classes cause the rejection of some known samples. Thus ALL* in T_{30} is low.

V. CONCLUSION

This paper considers a special fault diagnosis scenario where the label set of training data (source data) is a subset of the label set of test data (target data). For the SOSFD, EVT is applied to learned features to establish a rejection model. Accepted samples are classified according to SoftMax probabilities as normal. For the COSFD, weighted domain adversarial neural networks is proposed to obtain domain-invariant features on shared classes and separate samples of unknown classes in target domain from other samples before the process of open set recognition. Experimental results on the CWRU dataset validate that the proposed method has the ability to effectively detect samples from the unknown classes and ensure the accurate classification accuracy of samples from known classes simultaneously.
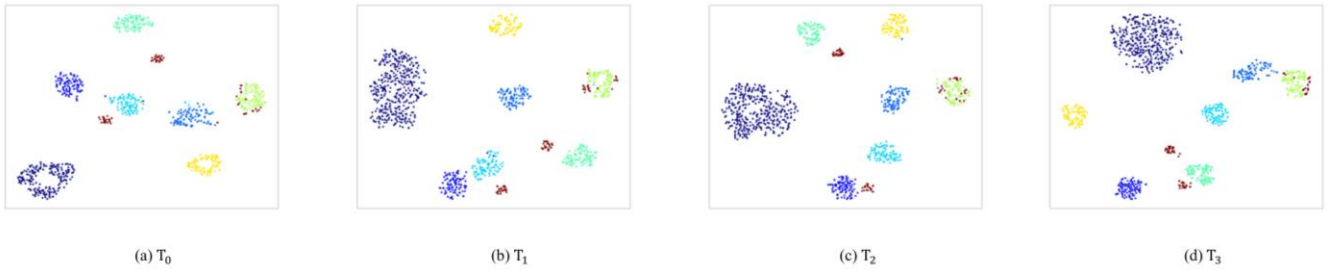
Fig. 3. Feature visualizations of training data and test data for shared-domain open set fault diagnosis. Red points represent unknown samples in test data. Points in other colors represent known samples in training data and test data.
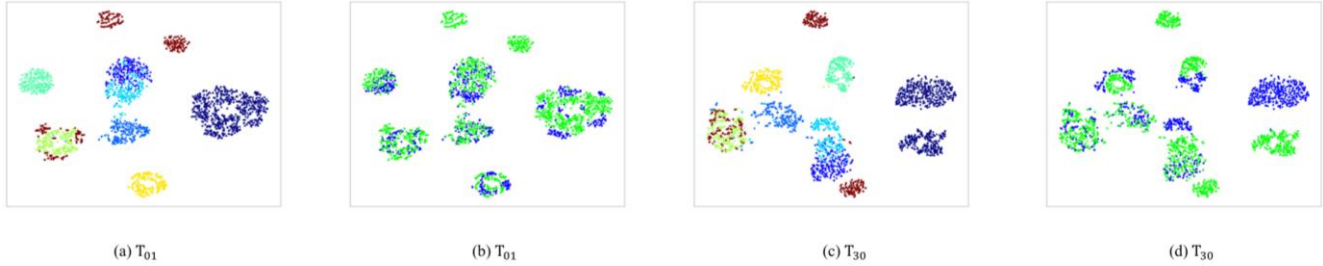
Fig. 4. Feature visualizations of source data and target data for cross-domain open set fault diagnosis. In (a) and (c), red points represent unknown samples in target data and points in other colors represent known samples in source data and target data. In (b) and (d), blue points represent source samples and green points represent target samples.

REFERENCES

[1] Zhao Z, Li T, Wu J, et al. Deep Learning Algorithms for Rotating Machinery Intelligent Diagnosis: An Open Source Benchmark Study[J]. ISA Transactions, 2020.

[2] Zhao Z, Zhang Q, Yu X, et al. Unsupervised Deep Transfer Learning for Intelligent Fault Diagnosis: An Open Source and Comparative Study[J]. arXiv preprint arXiv:1912.12528, 2019.

[3] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 35, no. 7, pp. 1757-1772, 2012.

[4] C. Geng, S. J. Huang, and S. Chen, "Recent Advances in Open Set Recognition: A Survey," *IEEE Trans Pattern Anal Mach Intell,* Mar 18 2020, doi: 10.1109/TPAMI.2020.2981604.

[5] A. Bendale, T. E. Boult, and Ieee, "Towards Open Set Deep Networks," in *2016 Ieee Conference on Computer Vision and Pattern Recognition*, (IEEE Conference on Computer Vision and Pattern Recognition. New York: Ieee, 2016, pp. 1563-1572.

[6] Z. Ge, S. Demyanov, Z. Chen, and R. Garnavi, "Generative openmax for multi-class open set classification," *arXiv preprint arXiv:1707.07418,* 2017.

[7] P. Oza and V. M. Patel, "C2AE: class conditioned auto-encoder for open-set recognition," (in English), *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Proceedings,* Conference Paper pp. 2302-2311, 2019 2019, doi: 10.1109/cvpr.2019.00241.

[8] R. Yoshihashi, W. Shen, R. Kawakami, Y. Shaodi, M. Iida, and T. Naemura, "Classification-Reconstruction Learning for Open-Set Recognition," (in English), *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Proceedings,* Conference Paper pp. 4011-4020, 2019 2019, doi: 10.1109/cvpr.2019.00414.

[9] Y. Tian, Z. L. Wang, L. P. Zhang, C. Lu, and J. Ma, "A subspace learning-based feature fusion and open-set fault diagnosis approach for machinery components," (in English), *Adv. Eng. Inform.,* Article vol. 36, pp. 194-206, Apr 2018, doi: 10.1016/j.aei.2018.04.006.

[10] M. A. Chao, B. T. Adey, and O. Fink, "Knowledge-Induced Learning with Adaptive Sampling Variational Autoencoders for Open Set Fault Diagnostics," *arXiv preprint arXiv:1912.12502,* 2019.

[11] P. P. Busto, J. Gall, and Ieee, "Open Set Domain Adaptation," in *2017 Ieee International Conference on Computer Vision*, (IEEE International Conference on Computer Vision. New York: Ieee, 2017, pp. 754-763.

[12] K. Saito, S. Yamamoto, Y. Ushiku, and T. Harada, "Open set domain adaptation by backpropagation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 153-168.

[13] H. Liu, Z. Cao, M. Long, J. Wang, and Q. Yang, "Separate to adapt: Open set domain adaptation via progressive separation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2927-2936.

[14] W. J. Scheirer, A. Rocha, R. J. Micheals, and T. E. Boult, "Meta-Recognition: The Theory and Practice of Recognition Score Analysis," *IEEE Trans Pattern Anal Mach Intell,* vol. 33, no. 8, pp. 1689-95, Aug 2011, doi: 10.1109/TPAMI.2011.54.

[15] M. Mundt, I. Pliushch, S. Majumder, and V. Ramesh, "Open Set Recognition Through Deep Neural Network Uncertainty: Does Out-of-Distribution Detection Require Generative Classifiers?," (in English), *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW),* Conference Paper pp. 753-757, 2019 2019, doi: 10.1109/iccvw.2019.00098.

[16] Y. Ganin and V. Lempitsky, "Unsupervised Domain Adaptation by Backpropagation," *arXiv: Machine Learning,* 2014.

[17] I. Goodfellow *et al.*, "Generative Adversarial Nets," in *neural information processing systems*, 2014, pp. 2672-2680.

[18] J. Zhang, Z. Ding, W. Li, and P. Ogunbona, "Importance Weighted Adversarial Nets for Partial Domain Adaptation," in *computer vision and pattern recognition*, 2018, pp. 8156-8164.

[19] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," Mechanical Systems and Signal Processing, vol. 64-65, pp. 100–131, 2015.

[20] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," *arXiv preprint arXiv:1409.7495,* 2014.

[21] L. V. Der Maaten and G. E. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research,* vol. 9, pp. 2579-2605, 2008.

A Novel Polarized Light Compass

Aided By GPS Module

Chengshuai Zhao
School of Instrument and Electronics
North University of China,
Taiyuan, China
1012679134@qq.com

Xiaojie Liu
School of Instrument and Electronics
North University of China,
Taiyuan, China
s1706092@st.nuc.edu.cn

Chong Shen
School of Instrument and Electronics
North University of China,
Taiyuan, China
shenchong@nuc.edu.cn

Jun Tang*
School of Instrument and Electronics
North University of China,
Taiyuan, China
tangjun16@126.com

Jun Liu
School of Instrument and Electronics
North University of China,
Taiyuan, China
liuj@nuc.edu.cn

Abstract—**This paper proposes and designs a novel bionic polarized light compass with the aid of the microarray polarized camera. The compass contains a small Global Positioning System module to assist orientation. The algorithm uses the Principal Component Analysis method to fit the solar meridian, and combines the astronomical calendar method to output the heading angle of the carrier in real time. The Root Mean Squared Error is 0.48 ° under dynamic conditions and 0.18 ° under rotation conditions, which has achieved good results in outdoor navigation experiments.**

Keywords—polarized light compass, bionic，orientation, navigation

I. INTRODUCTION

Currently, the most widely used navigation technology is based on GNSS and INS or integrated navigation of both. GNSS as a passive navigation method, the initial heading angle is provided by the dual antenna and then INS or IMU can therefore give the absolute heading angle in the geographic coordinate. However, in the special period when the satellite antenna signal is interfered, deceived or destroyed, orientation can only rely on inertial measurement units or inertial navigation system. But the errors from INS accumulated over time. There will be serious obstacles in the application.

The extraordinary ability of animals that can make use of the atmosphere polarization mode provides important enlightenment for autonomous robot navigation.In the process of propagating to the earth, the sunlight is scattered by various particles in the earth's atmosphere, and a relatively stable polarization state distribution is formed in the whole sky. Researchers have discovered that some insects in nature rely on their polarized vision to obtain polarized light in the sky and identify their flight heading according to their distribution patterns. For example, bees depend on their compound eyes to travel between the hive and the feeding point [1]. An ant in the Sahara desert, after a long journey in random path for food without olfactory failure and landmark information, uses the sky polarization mode to obtain direction information, and returns to the nest in the route of straight line approximately [2][3]. Compared with the traditional navigation methods such as global navigation satellite system, geomagnetic navigation and inertial navigation, bionic atmospheric polarized light navigation mainly relies on the sun position in the sky, which has obvious advantages of anti-interference, strong robustness and less error accumulation.

Although all types of imaging polarimeters exist, they may not qualify for the research on the imaging polarization navigation algorithm. Many research groups have studied polarized vision-based orientation methods. In 2015, a real-time imaging orientation determination system was designed and implemented by Lu Hao and Zhao Kaichun [4][5]. The algorithm based on Hough transform can accurately extract the sun meridian in the image according to the features of the solar meridian and the patterns of the polarized skylight.The pattern recognition algorithm in most polarization imaging algorithms involves several nonlinear calculations and that imposed a significant computation burden. In 2017, They optimized the calculation method, simplified the non-linear calculation to the linear calculation, and realized its function on the digital signal processor, which can meet the practical requirements of low calculation and high precision in the embedded system.

Besides, Kong, Xiang long [6] et al. proposed a method to improve the performance of the visual inertial navigation system (VINS) using a biologically-inspired polarized compass in 2016. They integrated the homemade polarized compass into the visual inertial navigation pipeline. The practical tests in different environments proved the effectiveness and feasibility of the method. In 2017, Han, Guo liang [7] and his research team proposed a polarized light compass composed of a charge-coupled device (CCD) camera, a pixelated polarizer array, and a wide-angle lens, and proposed a skylight polarization pattern based on a single-scatter Rayleigh model. In outdoor environment, the sensor can measure the skylight polarization pattern in real time. The standard deviation of the positioning error is 0.15 degrees. The experimental results proved that the polarized navigation sensor can be used for autonomous outdoor navigation. In the same year, Wang, Yu jie [8] and his team proposed a method for extracting heading information from the sky light polarization pattern, and for the first time successfully implemented a polarized light compass (PL-compass) in a heavily occluded foliage environment. The results show that the position error of the integrated navigation system can be greatly reduced with the help of the PL compass. The estimated trajectory tracks the ground truth very well. In the experiment, and the heading error is less than 1 degree.

---

*Jun Tang is the corresponding author. (e-mail: tangjun16@126.com).

This paper presents an integrated polarized light compass aided by GPS module, which can not only output the position parameters (including longitude and latitude) of the carrier in real time, but also output the attitude parameters of the carrier. We mainly studied its orientation method. The compass consists of a micro-array polarizing camera, a mini GPS module and a high-performance NVIDIA processor. Based on the inspiration of biological polarization vision, it provides stable heading information for single-antenna GPS. The structure of the system is simple and easy to install. After several outdoor experiments, the orientation accuracy is significantly better than GPS alone.

II. SYSTEM STRUCTURE OF POLARIZED LIGHT COMPASS

The system composition of the designed pixelated polarized light compass is mainly composed of a mini-GPS receiver, a NVIDIA processor, a CMOS camera with a pixelated polarizer array and a wide-angle lens.
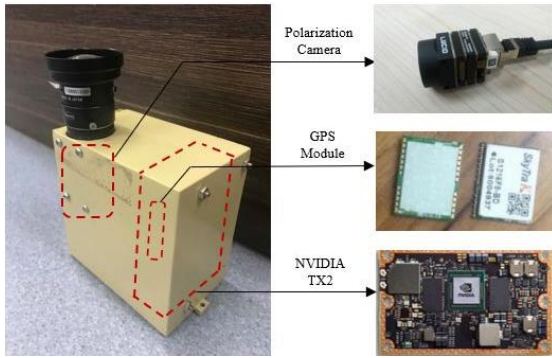
Fig. 1. System structure of parized light compass.

The CMOS polarization camera(Lucid,TRI050S-P/Q) is the most important sensor in the polarized light compass. It has a resolution of 2048×2448 .The focal length of the wide-angle lens is 3.5mm and the field of view of which is about $120° \times 120°$. What makes it special is the pixelated polarizer array. The design inspiration of the pixelated polarizer array comes from the compound eyes of insect. The compound eyes of insects such as sand ants and bees that use polarized light to navigate consist of countless small eyes. There is an area facing the sky in the compound eye, and its visual axis is upward, called Dorsal Rim Area (DRA). The dipole axe of the rhabdom pigment molecules in this area is parallel to the long axis of the microvilli. When the direction of the E-vector is parallel to the long axis of the microvilli, photons can be absorbed to the greatest extent. This special physiological structure makes DRA has a high polarization sensitivity [9].

Fig. 2. DRA and small eyes of insect.

Imitating the small eyes of insects, photodiode、micro-lenses and polarizer are integrated in each pixel. As is shown

in Fig. 3, the polarization sensor is composed of a 1024×1224 array of unit cells, each of which consists of a 2×2 array of polarizers on the surface of imaging units. The adjacent four polarizers in one unit cell pass linearly polarized light along different directions. The top left-hand and the bottom left-hand polarizers pass linearly polarized light oriented along the x-axis and at 45◦ counterclockwise to the x-axis respectively, while the polarization direction of the right two polarizers are perpendicular to that of the left two polarizers. The aperture of each micro-polarizer is 3.4 µm×3.4 µm, the size of which is exactly the same as the size of CMOS camera pixels. Compared with the ordinary high-pixel camera, in addition to the original picture, it can also obtain a polarization angle image, a polarization degree image, and four gray scale images of light passing through four different polarization angles. It connected to the NVIDIA processor through a Gigabit Ethernet cable to complete the image transmission. At the maximum resolution, the transmission rate can reach 8 times per second.
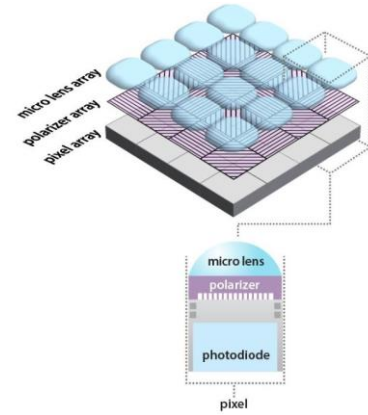
Fig. 3. Structure of polarization sensor.

Core component of the polarized light compass is a NVIDIA JETSON TX2 processor, which owns rich peripherals and interfaces for hardware, a stable Linux operating system, powerful computing capabilities, and high-quality open source function libraries for software, especially for computer vision processing. When it is connected with a monitor, we can regard it as a host or terminal with super functions. It can run with the same functions as PCs and can also access the Internet through a network cable or wirelessly. We calculate and fuse the images obtained by polarizing cameras and the data obtained from other sensors to make navigation data more efficient and accurate.

Satellite signals rely mainly on SkyTra-S1216F8-BD for processing in mini-GPS receiver. It interacts with TX2 through the serial port. It can provide relatively accurate UTC time as well as longitude, latitude and velocity. These data facilitate the real-time calculation of the solar height angle and solar azimuth.

III. ALGORITHM DESCRIPTION

According to the principle of Rayleigh scattering, the E-vector of scattered light is perpendicular to the surface formed by the incident light and the scattered light [4].The Rayleigh sky model can express the angle of the E-vector of a beam clockwise with respect to the local meridian as

$$tan\alpha = \frac{cos(hp)\,sin(hs) - sin(hp)\,cos(hs)\,cos(\varphi s - \varphi p)}{cos(hs)\,sin(\varphi s - \varphi p)} \qquad (1)$$

The x and y axes are parallel to the width and height axes of the image sensor respectively, the z axis is perpendicular to it. In Eq. (1), h_p and φ_p are the vertical angle and azimuth of the observation point, h_s and φ_s are the vertical angle and azimuth of the sun. As is shown in Fig. 4, when the observation direction is perpendicular to the zenith$(h_p = \frac{\pi}{2})$, the direction of the E-vector is parallel to the plane of the image sensor and perpendicular to the solar meridian.
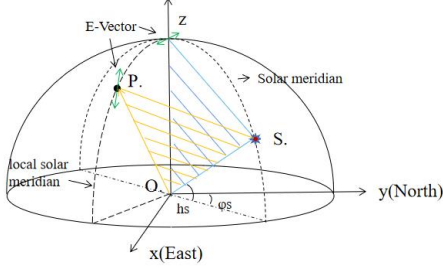
Fig. 4. E-vector of polarized light in the celestial sphere.

A. Procedures of the Algorithm

The algorithm in the polarized light compass is mainly divided into the following five steps as shown in Fig. 5.

1) Obtain the image of 0 °、45 °、90 °、135 ° channel as source image from polarization camera.

2) Calculate Stokes parameters and generate the full-sky degree of polarization and angle of polarization.

3) Coordinate transformation for recognition: Move the origin of the reference coordinate system of the polarization angle to the center of the image, and regenerate the image of angle of polarization.

4) Extract the symmetry axis of the image and obtain the angle between the solar meridian and the axis of the carrier.

5) Calculate the angle between the solar meridian and true north through the astronomical calendar using the local time and position.
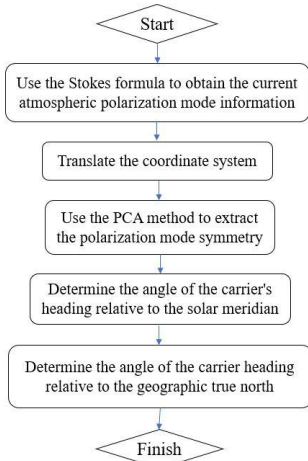
Fig. 5. Flowchart of the algorithm.

B. Obtain Image of The Full-sky Polarization Information

In order to measure the polarization mode of the sky, we use the Stokes vector to calculate the degree of polarization and angle of light in the sky. The Stokes vector is be marked as [I,Q,U,V], I represents the total intensity of a beam of light, Q represents the light intensity component of a beam of light

along the 0 degree polarization direction; U represents the light intensity component of a beam of light in a 45-degree polarization direction; V represents the circularly polarized light component, and generally speaking, V is negligible. (V = 0),The angle of polarization and degree of polarization of light are expressed as

$$AOP = \frac{1}{2}arctan\left(\frac{U}{Q}\right) \qquad (2)$$

$$DOP = \frac{\sqrt{Q^2+U^2}}{I} \qquad (3)$$

When the polarization direction of the linear polarizer is α with the optical axis of the system, easy to get the intensity of light with Mueller matrix

$$I_{out}(\alpha) = \frac{1}{2}(I_{in} + Q_{in}\cos 2\alpha + U_{in}\sin 2\alpha) \qquad (4)$$

I_{out} is the intensity of light in the α polarization direction of the linear polarizer. The detection method in this article is based on the relationship between the light intensity information and the Stokes parameter in the four polarization directions of 0 °、45 °、90 °、135 °, and the polarization information is obtained by detecting the brightness. Therefore, we need obtain four brightness images with polarization information in the same sky area $I_{0°}$、$I_{45°}$、$I_{90°}$、$I_{135°}$.From this we can get the Stokes parameter for each point in the area

$$I(x,y) = \frac{1}{2}[I_{0°}(x,y) + I_{45°}(x,y) + I_{90°}(x,y) + I_{135°}(x,y)] \qquad (5)$$

$$Q(x,y) = I_{0°}(x,y) - I_{90°}(x,y) \qquad (6)$$

$$U(x,y) = I_{45°}(x,y) - I_{135°}(x,y) \qquad (7)$$

In fact, using three light intensity images $I_{0°}, I_{45°}$, and $I_{90°}$ containing polarization information can already establish the relationship between the brightness information and the Stokes parameter

$$I(x,y) = I_{0°}(x,y) + I_{90°}(x,y) \qquad (8)$$

$$Q(x,y) = I_{90°}(x,y) - I_{0°}(x,y) \qquad (9)$$

$$U(x,y) = I_{0°}(x,y) + I_{90°}(x,y) - 2I_{45°}(x,y) \qquad (10)$$

C. Coordinate Transformation for Recognition

In order to obtain a single Rayleigh scattering E-vector image of the apex region, a reference transformation is performed on the image coordinates. First translate the coordinate system. Take the center point of the image as the principal point,

$$x' = x - \frac{w}{2}, y' = y - \frac{h}{2} \qquad (11)$$

$$h_p = arccot(\frac{\sqrt{x'^2+y'^2}}{f}), \quad \phi_p = arctan(\frac{y'}{x'}) \qquad (12)$$

According to pinhole camera model, to get the angle between the direction of the E-vector and the local solar meridian,
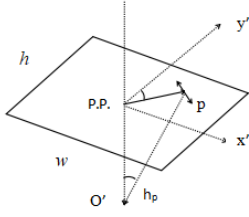
Fig. 6. Pinhole camera model.

taking the image center point as the origin of the coordinate system, the direction of the E-vector after the reference transformation is expressed as

$$AOPE(x', y') = AOP(x', y') - \phi_p(x', y') \quad (13)$$

It should be noted that the calculation results of the polarization direction calculated by this method need to have constraints：

$$AOP(x, y) = AOP(x, y) - \frac{\pi}{2}, \quad I_{0°} > I_{90°} \quad (14)$$

$$AOP(x, y) = AOP(x, y) + \frac{\pi}{2}, \quad I_{0°} \leq I_{90°} \cap I_{0°} + I_{90°} < 2I_{45°} \quad (15)$$

$$AOP(x, y) = AOP(x, y) - \frac{\pi}{2}, \quad I_{0°} \leq I_{90°} \cap I_{0°} + I_{90°} \geq 2I_{45°} \quad (16)$$

The degree of polarization and the symmetry of the polarization angle after the reference transformation can be observed by observation. Then we extract the feature points by setting a threshold of the degree of polarization and the polarization angle, and we fit an approximately straight solar meridian in the image. If the angle of the E-vector of a beam equals 90°, the projection of this beam must be on the solar meridian. In conclusion, E-vector of 90°、straight line and through the principal point are three features of the solar meridian.

D. Principal Component Analysis（PCA）

Then in order to further extract the accurate axis of symmetry in the image, the principal component analysis is performed in the algorithm. the symmetry detection method based on the principal component analysis transforms the solution of the direction of the symmetry axis into an eigenvalue decomposition problem of the covariance matrix. Specifically，supposed that set of points extracted according to the set threshold method expressed as
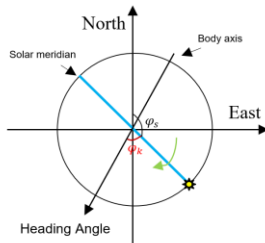
Fig. 7. Cartesian coordinates of atmospheric polarization mode.

$$P = [p1, p2, \cdots, \ pn]_{n \times 2} \quad (17)$$

Using these feature points as principal components, compute its covariance matrix $P^T P$ and its eigenvalue decomposition. The obtained two eigenvalues correspond to two eigenvectors,

which respectively represent the symmetry axis of the image and its vertical direction. We select eigenvectors with smaller eigenvalues to calculate the inclination of the solar meridian φ_k. Define clockwise as positive direction. The heading angle of the carrier is expressed as

$$\phi_z = \phi_s + \phi_k \quad (0 < \phi_z < 2\pi) \quad (18)$$

where ϕ_s is the absolute solar azimuth calculated using the astronomical calendar When the axis of the carrier body coincides with the meridian of the sun, it is recorded as

$$\phi_k = 0° \text{ or } \phi_k = 180° \quad (19)$$

When using it as a reference for navigation, it is clear that there is an ambiguity of 180 ° in the angle. Therefore, we defined the positive and negative directions of the carrier body axis and solar meridian. When installing the sensor on the carrier, we keep the 0° reference direction parallel to the positive direction of the carrier body axis. When the positive directions of the carrier body axis and solar meridian coincide, it is recorded as

$$\phi_k = 0°, \quad (20)$$

on the contrary

$$\phi_k = 180° \quad (21)$$

Because of its statistical characteristics, this method is not very accurate and can only be used as a rough extraction process of the symmetry axis. It needs to be optimized near this direction.

E. Astronomical Ephemeris Algorithm

Through the above method, we obtained the inclination of the solar meridian in the image coordinate system, denoted as ϕ_k. However, we can not get accurate heading angle of the carrier when if we not clear the absolute angle ϕ_s of solar meridian in the geographic coordinate system.

The information source of the astronomical ephemeris algorithm is the mini-GPS module of the system. Among them, year, month, day, hour, minute, second, longitude and latitude are necessary known quantities. In the astronomical ephemeris algorithm, we regard the sun as a point light source on the surface of the celestial sphere. The solar height angle is recorded as θ_s and the solar azimuth is recorded as ϕ_s. According to the astronomical ephemeris algorithm, we can get the sun position $s(r, \theta_s, \phi_s)$ when we have known the sun declination angle δ、the solar hour angle T and latitude of the observer l,

$$cos\theta_s = sin\delta sinl + cos\delta cosT, \quad 0 \leq \theta_s \leq \frac{\pi}{2} \quad (22)$$

$$cos\phi_s = \frac{sin\delta - cos\theta_s sinl}{sin\theta_s cosl}, \quad 0 \leq \phi_s \leq 2\pi \quad (23)$$

where the solar declination angle δ is a single-valued function of the solar day angle α.

$$\delta(\alpha) = 0.3723 + 23.2567sin\alpha + 0.1149sin2\alpha - 0.1712sin3\alpha - 0.758cos\alpha + 0.3656cos2\alpha + 0.0201cos3\alpha \quad (24)$$

Solar day angle α is expressed as

$$\alpha = \frac{2\pi(D-D_0)}{365.2422}, \ 0 < D < 366 \qquad (25)$$

$$D_0 = 79.6764 + 0.2422 \times (Y - 1985) - INT[\frac{Y-1985}{4}] \quad (26)$$

In the Eq. (26), D represents the accumulated days and Y represents the year. In order to find the true solar time and hour angle, we have to calculate local time S_d and jet lag E_t firstly,

$$S_d = S_O + \frac{F_O - [120° - (J_D + \frac{J_F}{60})] \times 4}{60} \qquad (27)$$

$$E_t(\alpha) = 0.0028 - 1.9857 sin\alpha + 9.9059 sin2\alpha - 7.0924 cos\alpha - 0.6882 cos2\alpha \qquad (28)$$

where S_O and F_O are Beijing hours and minutes, J_D and J_F are longitude and longitude minutes of observation point. The corrected true solar time S_t and hour angle T are expressed as

$$S_t = S_d + \frac{E_t}{60} \qquad (29)$$

$$T = (S_t - 12) \times 15° \qquad (30)$$

The absolute direction of the solar median in the geographic coordinate is clearly when the symmetry axis detection and astronomical ephemeris algorithm finished.

IV. EXPERIMENTS AND RESULTS

A. Static Experiment

In the afternoon of May 14, 2019, we set up a static experimental platform on the top of the Research building of North University of China （112.449451° E, 38.016014° N） and conducted experiments. For sake of simplicity, we supposed the heading angle as 90 ° at static experiment. we parallelized the camera's 0-degree polarization direction with the true south direction and 90-degree polarization direction with the true east direction. Then we obtain the original image, a polarization angle image, a polarization degree image, and four gray scale images of light passing through four different polarization angles in several times at about 4:00pm. The azimuth and height angle of the sun are 262.2854 ° and 40.254 ° . After that, the polarization angle images were processed on MATLAB. We set the threshold of polarization angle to 89.9 and the threshold of degree of polarization to 0.015. Then the solar meridian was extracted according to the principal component analysis algorithm. The correctness of the algorithm was experimentally verified. The experimental data obtained from measurements is shown in the table below.
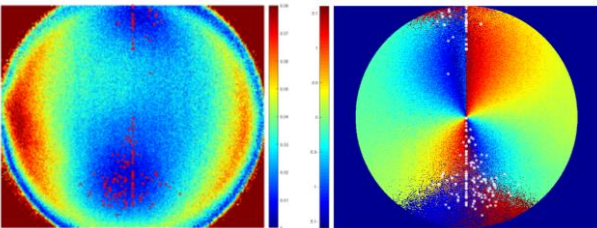
Fig. 8. Feature points in degree of polarization and angle of E-vector.

TABLE I. RESULTS AFTER PRINCIPAL COMPONENT ANALYSIS

Time	Eigenvectors		Eigen values	Symmetry axis inclination
15:57:54	-0.9966	-0.0825	0.0947	184.7308
	-0.0825	0.9966	2.2908	175.2692
16:04:37	-0.9968	-0.0801	0.0683	184.5958
	-0.0801	0.9968	1.5989	175.4042
16:05:55	-0.9961	-0.0887	0.0656	185.0886
	-0.0887	0.9961	2.2797	174.9114
16:06:54	-0.9985	-0.0546	0.0769	183.1302
	-0.0546	0.9985	2.1347	176.8698
16:09:04	-0.9956	-0.0941	0.0533	185.4015
	-0.0941	0.9956	2.5200	174.5985
16:10:24	-0.9970	-0.0772	0.0857	184.4265
	-0.0772	0.9970	1.7996	175.5735

TABLE II. RESULTS OF HEADING ANGLE

Time	Solar azimuth angle	Heading Angle
15:57:54	262.2854	87.0162
16:04:37	263.4554	88.0512
16:05:55	263.6787	88.7673
16:06:54	263.8526	86.9828
16:09:04	264.2155	89.6170
16:10:24	264.4410	88.8675

From the data we can see that it is difficult to fit the solar meridian by filtering the feature points only by setting the threshold and the direction is not necessarily very accurate, Using PCA method can reduce the dimension of the feature point data. In other words, it can find the symmetry axis of the image efficiently and get the slope of the solar meridian in the image.

B. Rotation Experiment

In addition, rotation experiments and vehicle experiment have been completed on the wheeled robots. In rotation experiments, we use a precise rotating stage and rotating controller (Zolix GT-111). In the experiment, the step of the rotating stage was 11° and the orientation was measured precisely. We recorded the heading angle and its errors. As is shown in the Fig. 11 and Fig. 12, the RMSE in the experiment is 0.1836 ° and the maximum error does not exceed 0.8 °.

Fig. 9. Rotation experimental arrangement.
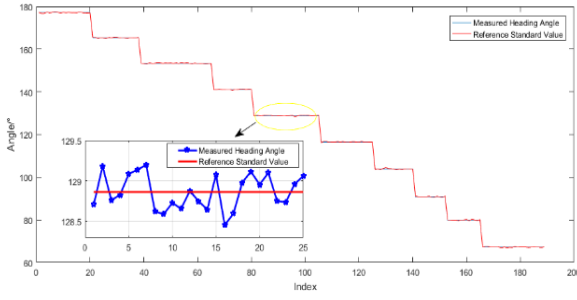
Fig. 10. Measurement azimuth of the rotation experiment.
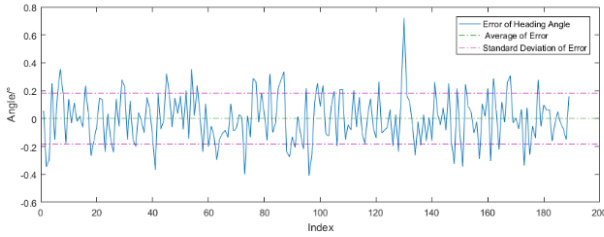
Fig. 11. Errors from azimuthal measurement of the rotation experiment.

C. Vehicle Experiment

In the vehicle experiment, we used our designed polarized light compass and high-precision GNSS/INS integrated navigation system for comparison, both of which installed on the same carrier and simultaneously measure the heading angle of the carrier. We regard the high-precision system as a reference. The driving route is set as a closed loop. The sampling frame rate of the PL-compass is 1Hz. The sampling rate of the GNSS/INS integrated navigation system is 50Hz. In order to compare the experimental data, we sampled the reference data. As is shown in the Fig. 15, the maximum error in the experiment does not exceed 1.5 °, the RMSE under dynamic conditions is 0.4863°.
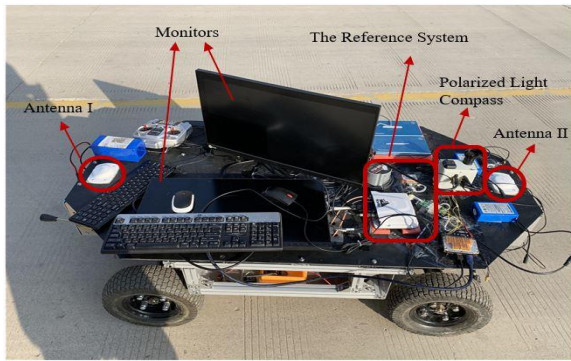
Fig. 12. Vehicle experimental arrangement.
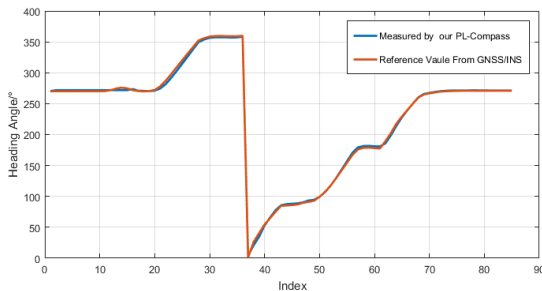
Fig. 13. azimuthal measurement of the vehicle experiment.
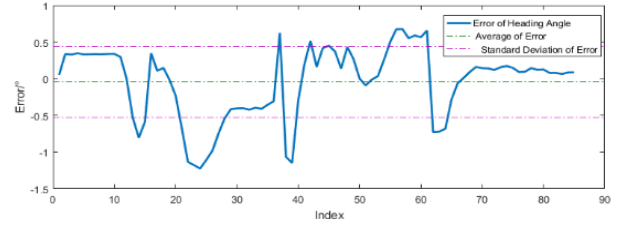
Fig. 14. Errors from azimuthal measurement of the vehicle experiment.

V. CONCLUSION

This paper analyzes the symmetry of the distribution of polarization angle and degree of polarization in the atmospheric polarization mode, proposes and designs a novel polarized light compass, which uses the PCA method to identify the solar meridian effectively. The GPS module combined with the ephemeris algorithm realizes the real-time orientation of the carrier, and solves the problem that the single antenna GPS module cannot be oriented accurately. The effectiveness of this method is verified through static experiments. The orientation accuracy have tested through rotation experiments and vehicle experiments. Under rotation conditions, the RMSE is 0.18 ° and the maximum error is 0.8 °. Under dynamic conditions, the RMSE is about 0.48 °, and the maximum error is less than 1.5 °. The test effect under static conditions is obviously better than that under dynamic conditions. In addition, because the PCA method is an algorithm based on statistical characteristics, the extraction method of the axis of symmetry still needs to be further optimized through filtering and denoising methods.

REFERENCES

[1] M. Sakura, R. Okada, and H. Aonuma, "Evidence for instantaneous e-vector detection in the honeybee using an associative learning paradigm," Proceedings of the Royal Society B, vol. 279, No. 1728, pp. 535–542, February 2012.

[2] J. Gao, L. Wang, M. Bo, and Z. Fan, "Information acquisition in desert ant navigation," International Journal of Information Acquisition, vol. 3, No. 1, pp. 33–43, February 2006.

[3] R. Wehner, "Desert ant navigation: how miniature brains solve complex tasks," Journal of Comparative Physiology A, vol. 189, No. 8, pp. 579–588, August 2003.

[4] H. Lu, K. Zhao, Z. You, and K. Huang, "Angle algorithm based on Hough Transform for imaging polarization navigation sensor," Opt. Express, vol. 23, No. 6, pp. 7248–7262, March 2015.

[5] H. Lu, K. Zhao, and Z. You, "Real-time polarization imaging algorithm for camera-based polarization navigation sensors," Applied optics, vol. 56, No. 1, pp. 3199–3205, April 2017.

[6] X. Kong, W. Wu, L. Zhang, X. He, Y. Wang. "Performance improvement of visual-inertial navigation system by using polarized light compass," The Industrial Robot, vol. 43, No. 6, pp. 588–595, October 2016.

[7] G. Han, X. Hu, J. Lian, "Design and calibration of a novel bio-inspired pixelated polarized light compass," Sensors, vol. 17, No. 11, p. 2623, November 2017.

[8] Y. Wang, X. Hu, "Polarized light compass-aided visual-inertial navigation under foliage environment," IEEE Sensor Journal, vol.17, NO.17, pp. 5646–5653, September 1 2017.

[9] J. Stalleicken, T. Labhart, H. Mouritsen, "Physiological characterization of the compound eye in monarch butterflies with focus on the dorsal rim area," Journal of Comparative Physiology A, vol. 192, No. 3, pp. 321–331, March 2006.

Structural Design and Simulation Analysis of Silicon Micro Triaxial Wheel-ring Gyroscope

Huiliang Cao*
Science and Technology on Electronic Test & Measurement Laboratory
North University of China
Tai Yuan, China
caohuiliang1986@126.com

Rui Zhao
Science and Technology on Electronic Test & Measurement Laboratory
North University of China
Tai Yuan, China
zhaorui@nuc.edu.cn

Qi Cai
Science and Technology on Electronic Test & Measurement Laboratory
North University of China
Tai Yuan, China
c18434363610@163.com

Yunbo Shi
Science and Technology on Electronic Test & Measurement Laboratory
North University of China
Tai Yuan, China
shiyunbo@nuc.edu.cn

Li Liu
Jinzhou 777 Microelectronics Co Ltd
Jin Zhou, China
327490545@qq.com

Abstract—**Micromechanical gyroscope is one of the important applications of attitude control platform and inertial navigation system. However, domestic gyroscope research and development is not highly integrated at present, and the main research focuses on the design of micro-gyroscope with a single detection axis (especially z-axis detection). In order to realize the monolithic triaxial gyro which can be used in the field of intelligent ammunition, a wheel ring gyro structure is designed. The gyro structure was modeled on the SolidWorks software, and then modal analysis, transient impact response analysis and harmonic response analysis were carried out in the finite element analysis software ANSYS respectively. According to the natural frequency of each mode, the structure size of the gyroscope is optimized. Through the simulation results, it is observed that the frequency split between the working modes is relatively small, and the structure has high mechanical sensitivity. The transient impact simulation of the gyro structure shows that the maximum stress of the sensitive structure under the impact of 10000g@5ms is 394.34Mpa, which proves that the structure has good impact resistance.**

Keywords—Three-axis gyroscope, MEMS, Simulation

I. INTRODUCTION

Integrating the sensor on the traditional artillery ammunition fuze can realize the intelligence of the ammunition. According to the different physical quantities that the sensor is sensitive to, the intelligent ammunition can achieve different military purposes. The three-axis (X, Y, Z) gyroscope can provide a three-dimensional attitude for the ammunition, which is a necessary condition for inertial guidance [1]-[4]. Therefore, the development of a gyroscope that can be used for smart ammunition is the top priority for smart ammunition to achieve inertial guidance.

At present, gyroscopes used to measure angular rate information of three axes at the same time adopt a gyroscope integrates three single-axis gyroscopes orthogonally arranged and a three-axis gyroscope with a monolithic structure. However, the integrated three-axis gyroscope with orthogonal placement [5][6] has a low degree of integration, large assembly errors, and a large system. The monolithic single-structure three-axis gyroscope[7]-[12] has complex structure, difficult processing, and serious coupling between modes. Therefore, these two types of gyroscopes are not suitable for the three-dimensional attitude detection of ammunition.

A new type of silicon micro three-axis wheel ring gyroscope based on MEMS technology is proposed in this paper, which uses a single-chip multi-structure gyroscope array method to integrate two gyroscope structures on a single chip. The gyroscope has simple structure, easy processing, simple decoupling, small size and easy protection. It can effectively provide a three-dimensional posture for the ammunition and realize the intelligentization of the ammunition. The vibration characteristics and transient shock response analysis of the resonant structure of the gyroscope have been carried out, which proves that the gyroscope has good vibration characteristics and shock resistance.

II. STRUCTURE DESIGN OF SILICON MICRO TRIAXIAL WHEEL-RING TYPE GYROSCOPE

Taking into account the size of the projectile and the high overload environment of the projectile during launch, the designed gyroscope structure should have the characteristics of small size and good impact resistance. The gyro structure proposed in this paper is realized by a single-chip three-axis wheel-ring gyro structure array. The two structures are realized by nesting to further reduce the structure volume. Figure 1 is a schematic diagram of the gyroscope structure. The wheel structure is completely nested in In the middle of the ring structure, the two structures work independently without affecting each other.

The wheel structure is from the inside to the outside by the Y-axis detection frame (inner cylindrical frame, XOY out-of-plane motion, torsion about the OX axis, used to detect the Y-axis input angular rate), Y-driven flexible joint (used to support Inner frame), wheel structure drive frame (medium cylindrical shell frame, connected with four fully symmetrical anchor points through four support beams to support the overall structure), X-driven flexible joint (used to support the outer frame) And X-axis detection frame (outer cylindrical shell frame, doing XOY out-of-plane motion, twisting around OY axis, used to detect X-axis input angular rate).

The structure of the ring-shaped solid wave gyro, as shown in Figure 1, is supported by eight symmetrical anchor points and S-shaped folding beams. The ring-shaped resonant structure in the middle is supported by the symmetrical structure. The working principle of the antinode, the structure has good impact resistance.
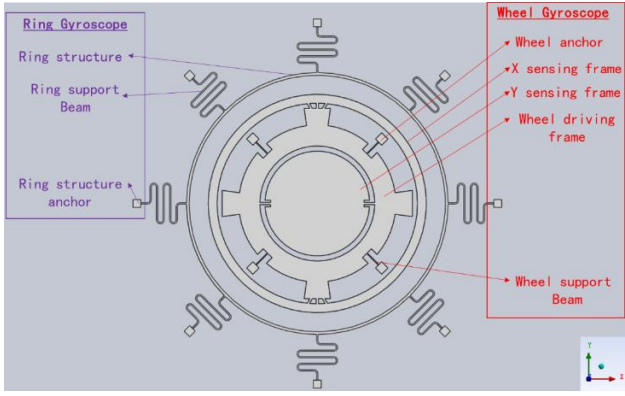
Fig 1. Schematic diagram of the monolithic three-axis ring type gyro structure

III. ANALYSIS OF VIBRATION CHARACTERISTICS

A. Modal Analysis

Through modal analysis, we can obtain the frequency and vibration characteristics of each mode of the gyroscope structure. First, use SolidWorks software to modeling the gyro structure, after that the finite element analysis software ANSYS is used to perform modal analysis on the gyroscope structure and obtain the modal information of each order of the gyroscope structure.. The first ten modes are shown in Figure 2. Table 1 lists the vibration characteristics and resonance frequency of each mode.
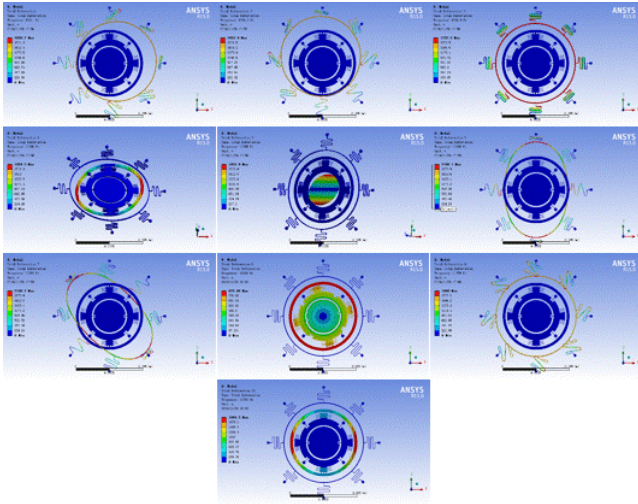
Fig 2. The first ten working modes of the ring gyro structure

In order to reduce the interference between modes, must increase the difference between the working mode frequency and the interference mode frequency. In order to improve the mechanical sensitivity of the gyro, the frequency split between the driving mode and the detection mode should be small but not exactly the same, because when they are exactly the same, the pulse width will be reduced [13]. It can be seen from the figures and tables that the working mode of the gyro is relatively concentrated, and the resonance frequency is far away from the resonance frequency of the interference mode, which avoids the influence of the interference mode on the operation of the gyro. The frequency split between the Z-axis drive and detection modes of the gyro is small (1 Hz), and the frequency split between the detection mode and drive mode of the wheel structure gyro X and Y axes is 160 Hz and 121 Hz, respectively, which greatly improves the mechanical sensitivity of the gyro. It can be seen that the gyro structure

can be far away from the influence of environmental vibration, and it has high mechanical sensitivity while ensuring the pulse width, so as to obtain higher anti-high overload performance.

TABLE I. The First Ten Working Modes Of The Three-axis Gyroscope

order	Form of exercise	Modal natural frequency	Remarks
First order	Ring structure translation in XOY plane（n=1）	9564.0Hz	Interference mode
Second order	Ring structure translation in XOY plane （n=1）	9566.3Hz	Interference mode
Third order	Ring structure translation in Z direction	9666.9Hz	Interference mode
Fourth order	The outer frame of the wheel structure is twisted around the OY axis	10166.0Hz	X-axis detection mode
Fifth order	The inner frame of the wheel structure is twisted around the OX axis	10205.0Hz	Y-axis detection mode
Sixth order	Four-node anti-vibration in XOY plane of ring structure （n=2）	10320.0Hz	Z axis drive mode
Seventh order	Four-node anti-vibration in XOY plane of ring structure （n=2）	10321.0Hz	Z-axis detection mode
Eighth order	The inner and outer frames of the wheel structure are twisted around the OZ axis	10326.0Hz	X/Y axis drive mode
Ninth order	Ring structure rotates around OZ axis	11759.0Hz	Interference mode
Tenth order	The outer frame of the wheel structure translates along the OZ axis	11763.0Hz	Interference mode
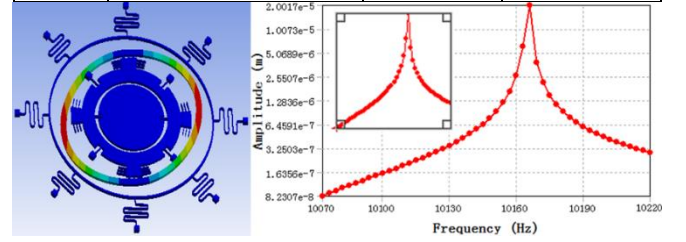
Fig 3. X-axis detection modal harmonic response analysis
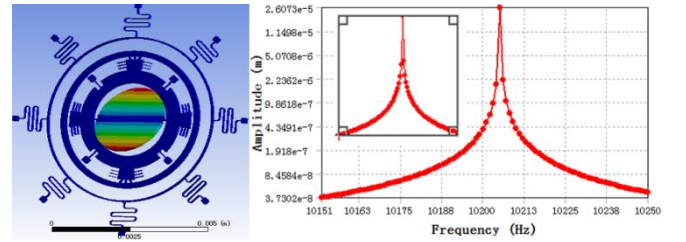
Fig 4. Y-axis detection modal harmonic response analysis
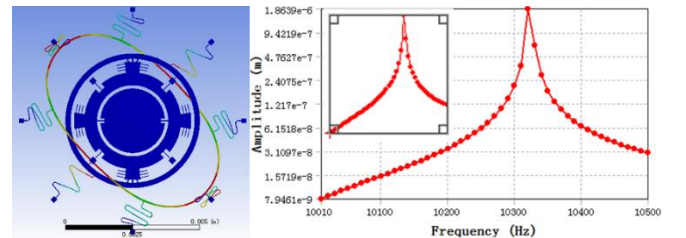
Fig 5. Z-axis detection modal harmonic response analysis

B. Harmonic Response Analysis

In order to determine the steady-state response of the structure subject to the time-constant harmonic change load, the harmonic response analysis of the structure. Through harmonic response analysis, the amplitude-frequency

response curve of the structure can be obtained, and we can observe at which frequency the resonance peak appears on the resonator, and obtain the magnitude of the peak.

On the basis of modal simulation, analyze the harmonic response of the structure for each working mode: apply a harmonic force with an amplitude of 1μN to the driving position of the structure, observe the displacement of the structure detection mode, and obtain the wheel structure X-axis detection mode (fourth-order mode, 10166.0 Hz, as shown in Figure 3), Y-axis detection mode (fifth-order mode, 10205.0 Hz, as shown in Figure 4) and circular Z-axis detection mode (The seventh-order mode, 10321.0 Hz, as shown in Figure 5). The simulation results show that there are only unique values in the resonant frequency range of each mode, and the X/Y axis detection displacement is large (indicating that the structure has better mechanical sensitivity), and the Z axis detection displacement is small (reflecting the higher sensitivity of the Z axis). Low, large range).

C. Transient Shock Response Analysis

When the gyroscope is used in smart ammunition, it needs to withstand large transient acceleration shocks during the launch of the ammunition. The impact time of this kind of impact load is short, the peak value is high, and the change is fast. Therefore, the gyroscope generates a large impact stress and is easy to damage the resonance structure. This article uses half-sine pulses to simulate the actual overload of the gyroscope, and uses the ANSYS transient analysis module for simulation.

A half-sine periodic shock load of 10 000g @ 5ms is applied in the z-axis direction of the gyro resonance structure. The result is shown in Figure 6. The maximum stress of the structure is 394.34MPa, which is much lower than the allowable stress (790MPa) of the silicon structure. The maximum stress point is distributed on the support beam of the pulley structure, close to the anchor point. It shows that the structure has good impact resistance and can adapt to the high overload environment when the ammunition is launched.
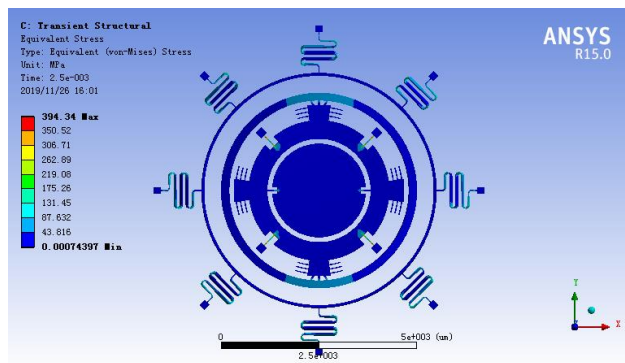
Fig 6. Stress analysis of gyro structure under 10 000g impact

IV. CONCLUSION

In order to realize the intelligent three-axis gyroscope used in ammunition, a new type of monolithic three-axis MEMS gyroscope structure based on the wheel-ring form is proposed. The gyro structure was modeled on the SolidWorks software, and the mode, harmonic response and transient impact response were analyzed respectively in the finite element analysis software ANSYS. Simulation analysis results show that the frequency split between the working mode and the interference mode of the three-axis gyroscope is large, and the frequency split between the drive mode and the detection mode is small, which can stay away from the disturbance of environmental vibration as well as maintain a high sensitivity . Under the transient impact of 10000g@10ms, the maximum stress of gyro structure is far less than the limit allowable stress of silicon, which is 394.3mpa, proving that gyro structure has a good performance in impact resistance.

REFERENCES

[1] Zhao Y, Liu Y et al. Summary of Development of Ballistic Correction Fuze [J]. Journal of Detection and Control (5):5.

[2] Shkel A M . Precision navigation and timing enabled by microtechnology: Are we there yet?[C]// Micro-& Nanotechnology Sensors, Systems, & Applications III. Micro- and Nanotechnology Sensors, Systems, and Applications III, 2011.

[3] Habibi S , Cooper S J, Stauffer J M , et al. Gun hard inertial measurement unit based on MEMS capacitive accelerometer and rate sensor[C]// Position, Location and Navigation Symposium, 2008 IEEE/ION. IEEE, 2008.

[4] Lutwak R. Micro-technology for positioning, navigation, and timing towards PNT everywhere and always[C]// 2014 International Symposium on Inertial Sensors and Systems (ISISS). IEEE, 2014..

[5] Yoon S J, Park U. Design and analysis of MEMS vibrating ring gyroscope considering high-g shock reliability[J]. The Transactions of The Korean Institute of Electrical Engineers, 2015, 64(10): 1440-1447.

[6] Yan W, Leng X, Wang Z, et al. High precision tri-axial MEMS gyroscope module based on redundant implementation and sensor fusion[C]//2016 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). IEEE, 2016: 376-379.

[7] Iqbal F, Din H, Lee B. Single Drive Multi-Axis Gyroscope with High Dynamic Range, High Linearity and Wide Bandwidth [J]. Micromachines, 2019, 10(6): 410.

[8] Efimovskaya A, Yang Y, Ng E, et al. Compact roll-pitch-yaw gyroscope implemented in wafer-level epitaxial silicon encapsulation process[C]//2017 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL). IEEE, 2017: 1-2.

[9] Wisher S, Shao P, Norouzpour-Shirazi A, et al. A high-frequency epitaxially encapsulated single-drive quad-mass tri-axial resonant tuning fork gyroscope[C]//2016 IEEE 29th International Conference on Micro Electro Mechanical Systems (MEMS). IEEE, 2016: 930-933.

[10] Xia D, Xu L. Coupling mechanism analysis and fabrication of triaxial gyroscopes in monolithic MIMU [J]. Micromachines, 2017, 8(10): 310.

[11] Giomi E, Fanucci L, Rocchi A. Analog-CMDA based interfaces for MEMS gyroscopes [J]. Microelectronics Journal, 2014, 45(1): 78-88.

[12] Jeon Y, Kwon H, Kim H C, et al. Design and development of a 3-axis micro gyroscope with vibratory ring springs [J]. Procedia Engineering, 2014, 87: 975-978.

[13] Kou Z, Liu J, Cao H, et al. Investigation, modeling, and experiment of an MEMS S-springs vibrating ring gyroscope[J]. Journal of Micro/ Nanolithography Mems & Moems, 2018, 17(1):1.

A Novel State Estimation Algorithm Based on Variational Bayesians Method Applied to AUV

Haoqian Huang*
School of Energy and Electrical Engineering
Hohai University
Nanjing, China
hqhuang@hhu.edu.cn

Jiacheng Tang
School of Energy and Electrical Engineering
Hohai University
Nanjing, China
191306060010@hhu.edu.cn

Chao Wang
School of Energy and Electrical Engineering
Hohai University
Nanjing, China
wangchao@hhu.edu.cn

Abstract—The MEMS-grade inertial sensors are mainly used to provide navigation information because the sensors loaded in autonomous underwater vehicle (AUV) is very stringent. However, the marine environment inevitably leads to larger errors for MEMS sensors due to its complexity and change. The measurement noise, playing an important role in state estimation with high accuracy, is hard to be obtained. To deal with the problem mentioned above, a new algorithm which fuses variational Bayesians into nonlinear filtering is proposed. The proposed variational Bayesians state estimation (VBSE) algorithm not only can obtain the accuracy measurement noise, but also the dimension of the observation vector is augmented to obtain the position information at the discrete time. The real underwater experiments are performed and experiment results give a draw that the proposed VBSE performs in aspect of positioning accuracy of AUV and robustness than the traditional algorithms.

Keywords—Measurement loss, state estimation, underwater environment, variational Bayesian

I. INTRODUCTION

The autonomous underwater vehicles (AUVs) play an important role in underwater detection, surveying and underwater reconnaissance in the ocean [1]. The more accurate navigation and positioning methods for AUV are the necessary conditions to obtain effective information [2]. The navigation and positioning methods with high accuracy for AUV are the necessary conditions for acquiring the effective information, meanwhile, and these methods also are the core technologies for estimating the positions of AUV.

The underwater ambient is complicated and changeable because of all sorts of noise interferences, so it is difficult to modeling accurately for the underwater environment [3]. Besides, a number of underwater navigation methods by using electromagnetic transmitting cannot be applied underwater because of the signal decays rapidly underwater [4]. Meanwhile, the positioning error of MEMS system accumulates as time goes on [5]. The shortcomings mentioned above will cause the increase of the measurement error, which has negative effects on the positioning for an AUV [6]. Thus, the performance of estimation approach is crucial to work usually for an AUV [7].

The cubature Kalman filter (CKF) employs the third degree spherical radial rule to calculate some cubature points which are used to calculate nonlinear integral. It is an useful technique in the case of nonlinearity. This technique has been widely applied in many applications including target tracking, the cooperative localization of AUV, control and so on [8].Moreover, the measurement noise with high accuracy is difficult to be described for underwater integrated navigation system because the sensors performs

poor under the condition of the changeable environment [9]. The variational Bayesians (VB) method has the merit of simplicity, and the each iteration of the VB approach increases the approximation degree [10].

The paper proposes a variational Bayesian state estimation (VBSE) algorithm , integrating the VB approach with the CKF to obtain navigation information with higher accuracy for a long period of working. The proposed VBSE has a better robustness and higher determination accuracy.

The rest of the paper is given below, the new proposed algorithm is described in Section II. The simulations show a remarkable improvement in the performance of the VBSE among those algorithms in next Section. Section IV summarizes main conclusions.

II. THE PROPOSED VARIATIONAL BAYESIANS STATE ESTIMATION ALGORITHM

The process function and measurement function are denoted in (1)-(2):

$$d\mathbf{X}(t) = \mathbf{F}(t)\mathbf{X}(t)dt + \omega(t)dt \tag{1}$$

$$\mathbf{Z}(t_k) = [r_k, \theta_k]^T = h[\mathbf{X}(t_k)] + v(t_k) \tag{2}$$

where $\mathbf{X}(t) = [\mathbb{X}(t), v_x(t), \mathbb{Y}(t), v_y(t)]^T$ is the state vector at the time t; $\mathbb{X}(t)$ and $\mathbb{Y}(t)$ are the position of the AUV at the time t in the x and y directions, respectively; $v_x(t)$ and $v_y(t)$ are the velocity of the AUV at the time t in the x and y directions, respectively. r_k means the distance at the time k measured horizontally from the measurement point to the AUV; θ_k means the angle of the AUV relatives to measurement point at the time k. The position of the measurement point is $(\mathbb{X}_0, \mathbb{Y}_0)$. $v(t_k)$ and $\omega(t)$ are measurement noise and process noise, respectively.

The model of the AUV is denoted in Fig.1.
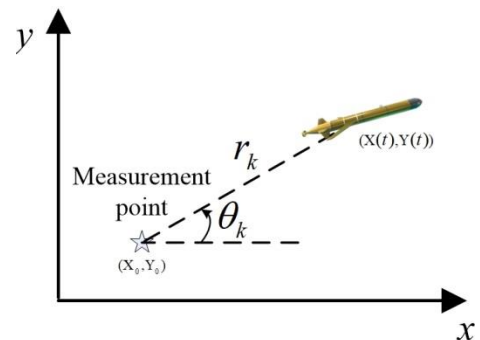
Fig. 1. The model of the AUV

The process function and measurement function are described as follows:

$$\mathbf{F}(t)=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$h[\mathbf{X}(t_k)]=\begin{bmatrix} \sqrt{[\mathrm{X}(t)-\mathrm{X}_0]^2+[\mathrm{Y}(t)-\mathrm{Y}_0]^2} \\ \arctan\left(\dfrac{[\mathrm{Y}(t)-\mathrm{Y}_0]}{[\mathrm{X}(t)-\mathrm{X}_0]}\right) \end{bmatrix} \tag{3}$$

To obtain more accuracy information of localization, the measurement equation can be rewritten by the state augmentation method:

$$\overline{\mathbf{Z}}(t_k)=[\mathbf{Z}^T(t_k) \quad \mathbf{L}^T(t_k)]^T \tag{4}$$

where $\mathbf{L}(t_k)=[\mathrm{X}(t_k),\mathrm{Y}(t_k)]^T$ means the localization information at the time k. Hence, the function $h(\bullet)$ needs to be expended as:

$$\overline{h}[\mathbf{X}(t_k)] \triangleq \begin{bmatrix} h[\mathbf{X}(t_k)] \\ \mathbf{H}_k\mathbf{X}(t_k) \end{bmatrix} \tag{5}$$

where $\mathbf{H}_k=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$. The joint posterior probability density function (PDF) $\mathrm{P}[\mathbf{X}(t_k),\mathbf{R}(t_k),\mathbf{P}(t_k^+)|\overline{\mathbf{Z}}_{1k}]$ is considered to estimate the positions and velocities of AUV in the two directions. $\overline{\mathbf{Z}}_{1k} \triangleq [\overline{\mathbf{Z}}(t_1),\overline{\mathbf{Z}}(t_2),\cdots,\overline{\mathbf{Z}}(t_k)]$ means all measurement information acquired from the sensors from time 1 to k; $\mathbf{P}(t_k^+)$ denotes the predicted error covariance matrix at the discrete time k; $\overline{\mathbf{Z}}(t_k)$ is the complete measurement information at discrete time k, but it is inaccurate to assume $\mathrm{P}[\mathbf{X}(t_k),\mathbf{R}(t_k),\mathbf{P}(t_k^+)|\overline{\mathbf{Z}}_{1k}]$ submit Gaussian distribution and the PDF is unknown. Hence, the VB approach is used to find some independent distribution $q(\bullet)$ to approximate $\mathrm{P}[\mathbf{X}(t_k),\mathbf{R}(t_k),\mathbf{P}(t_k^+)|\overline{\mathbf{Z}}_{1k}]$:

$$\mathrm{P}[\mathbf{X}(t_k),\mathbf{R}(t_k),\mathbf{P}(t_k^+)|\overline{\mathbf{Z}}_{1k}] \approx q[\mathbf{X}(t_k)]q[\mathbf{R}(t_k)]q[\mathbf{P}(t_k^+)] \tag{6}$$

$\mathbf{R}(t_k)$ and $\mathbf{P}(t_k^+)$ are chosen as inverse Wishart distribution.

Set $q[\mathbf{X}(t_k)]=\mathrm{N}(\mathbf{X}(t_k);\mathbf{X}(t_k^+),\mathbf{P}(t_k^+))$,

$q[\mathbf{R}(t_k)]=\mathrm{IW}(\mathbf{R}(t_k);u(t_k^+),\mathbf{U}(t_k^+))$

and $q[\mathbf{P}(t_k^+)]=\mathrm{IW}(\mathbf{P}(t_k^+);t(t_k^+),\mathbf{T}(t_k^+))$.

where $\mathrm{N}(\bullet;\mu,\sigma^2)$ means the variable follows Gaussian distribution with μ mean and σ^2 covariance; $\mathrm{IW}(\bullet;n,\mathbf{V})$ means that the variable follows inverse Wishart distribution with n degree of freedom (DOF) and $\mathbf{V}$ inverse scale matrix. Therefore, the core of the problem is to acquire the optimal results of state vector $\mathbf{X}(t_k)$, measurement matrix $\mathbf{R}(t_k)$ and the predicted error covariance matrix $\mathbf{P}(t_k^+)$. Let $\Theta=\{\mathbf{X}(t_k),\mathbf{R}(t_k),\mathbf{P}(t_k^+)\}$, and the optimal solutions are defined:

$$\log q(\xi)=\mathrm{E}_{\Theta^{-\xi}}[\log \mathrm{P}(\Theta,\overline{\mathbf{Z}}_{1k})]+C_\xi \tag{7}$$

where $\mathrm{E}_{\Theta^{-\xi}}[\bullet]$ means derivate the expectation of $\Theta^{-\xi}$, which means all the elements in Θ expect for ξ; C_ξ is a constant about ξ. Set $\xi=\mathbf{P}(t_k^+)$ and $\mathbf{R}(t_k)$, and the optimal solutions can be derived:

$$\begin{cases} \xi=\mathbf{P}(t_k^+) \\ \log q^{(l+1)}[\mathbf{P}(t_k^+)]=-0.5[n_\mathbf{X}+t(t_k^+)+2]\log|\mathbf{P}(t_k^+)| \\ \quad -0.5tr\{[\mathbf{A}^{(l)}(t_k)+\mathbf{T}(t_k^+)]\mathbf{P}^{-1}(t_k^+)\}+C_\mathbf{P} \\ \mathbf{A}^{(l)}(t_k)=\mathbf{P}(t_k)+[\mathbf{X}^{(l)}(t_k)-\mathbf{X}(t_k^+)][\mathbf{X}^{(l)}(t_k)-\mathbf{X}(t_k^+)]^T \end{cases} \tag{8}$$

$$\begin{cases} \xi=\mathbf{R}(t_k) \\ \log q^{(l+1)}[\mathbf{R}(t_k)]=-0.5[n_\mathbf{Z}+u(t_k^+)+2]\log|\mathbf{R}(t_k)| \\ \quad -0.5tr\{[\mathbf{B}^{(l)}(t_k)+\mathbf{U}(t_k^+)]\mathbf{R}^{-1}(t_k)\}+C_\mathbf{R} \\ \mathbf{B}^{(l)}(t_k)=\mathrm{E}^{(l)}\{\{\overline{\mathbf{Z}}(t_k)-\overline{h}[\mathbf{X}(t_k)]\}\{\overline{\mathbf{Z}}(t_k)-\overline{h}[\mathbf{X}(t_k)]\}^T\} \end{cases} \tag{9}$$

where $n_\mathbf{X}$ means the dimension of state vector; $n_\mathbf{Z}$ means the dimension of output vector augmented. The results from (8)-(9) are derived:

$$\begin{cases} q^{(l+1)}[\mathbf{P}(t_k^+)]=\mathrm{IW}\left(\mathbf{P}(t_k^+);t^{(l+1)}(t_k),\mathbf{T}^{(l+1)}(t_k)\right) \\ t^{(l+1)}(t_k)=t^{(l)}(t_k^+)+1 \\ \mathbf{T}^{(l+1)}(t_k)=\mathbf{A}^{(l)}(t_k)+\mathbf{T}^{(l)}(t_k^+) \end{cases} \tag{10a}$$

$$\begin{cases} q^{(l+1)}[\mathbf{R}(t_k)]=\mathrm{IW}(\mathbf{R}(t_k);u^{(l+1)}(t_k),\mathbf{U}^{(l+1)}(t_k)) \\ u^{(l+1)}(t_k)=u^{(l)}(t_k^+)+1 \\ \mathbf{U}^{(l+1)}(t_k)=\mathbf{B}^{(l)}(t_k)+\mathbf{U}^{(l)}(t_k^+) \end{cases} \tag{10b}$$

$$\mathbf{B}^{(l)}(t_k)=\mathrm{E}^{(l)}\left\{ \begin{bmatrix} \mathbf{Z}(t_k)-h[\mathbf{X}(t_k)] \\ \mathbf{L}(t_k)-\mathbf{H}_k\mathbf{X}(t_k) \end{bmatrix} \times \begin{bmatrix} \{\mathbf{Z}(t_k)-h[\mathbf{X}(t_k)]\}^T & [\mathbf{L}(t_k)-\mathbf{H}_k\mathbf{X}(t_k)]^T \end{bmatrix} \right\} \tag{11}$$

The cubature points are calculated below:

$$\chi_i^{(l)}=\mathbf{X}(t_k)+\sqrt{\mathbf{P}(t_k^+)}\delta_i \quad i=1,2,\dots 2n \tag{12}$$

where δ_i is ith column of matrix $\left[\sqrt{n}\mathbf{I}_{n\times n},-\sqrt{n}\mathbf{I}_{n\times n}\right]$, $\mathbf{I}_{n\times n}$ is n identity matrix.

When $\xi=\mathbf{X}(t_k)$, the logarithm to the $q[\mathbf{X}(t_k)]$ is updated:

$$\begin{aligned} \log q[\mathbf{X}(t_k)]=&-0.5(\overline{\mathbf{Z}}(t_k)-\overline{h}[\mathbf{X}(t_k)])^T\mathrm{E}^{(l+1)}[\mathbf{R}^{-1}(t_k)] \\ &\times(\overline{\mathbf{Z}}(t_k)-\overline{h}[\mathbf{X}(t_k)])-0.5[\mathbf{X}(t_k)-\mathbf{X}(t_k^+)]^T \\ &\times\mathrm{E}^{(l+1)}[\mathbf{P}^{-1}(t_k^+)][\mathbf{X}(t_k)-\mathbf{X}(t_k^+)]+C_\mathbf{X} \end{aligned} \tag{13}$$

where $\mathrm{E}^{(l+1)}[\mathbf{R}^{-1}(t_k)]=[u^{(l+1)}(t_k)-n_\mathbf{Z}-1]\mathbf{U}^{-(l+1)}(t_k)$;
$\mathrm{E}^{(l+1)}[\mathbf{P}^{-1}(t_k^+)]=[t^{(l+1)}(t_k)-n_\mathbf{X}-1]\mathbf{T}^{-(l+1)}(t_k)$.

$\mathbf{P}^{(l+1)}(t_k^+)$ and $\mathbf{R}^{(l+1)}(t_k)$ are calculated:

$$\begin{aligned} \mathbf{P}^{(l+1)}(t_k^+)&=\{\mathrm{E}^{(l+1)}[\mathbf{P}^{-1}(t_k^+)]\}^{-1} \\ \mathbf{R}^{(l+1)}(t_k)&=\{\mathrm{E}^{(l+1)}[\mathbf{R}^{-1}(t_k)]\}^{-1} \end{aligned} \tag{14}$$

Define $\mathrm{P}^{(l+1)}(\mathbf{X}(t_k)|\overline{\mathbf{Z}}_{1:k})=\mathrm{N}(\mathbf{X}(t_k);\mathbf{X}^{(l+1)}(t_k^+),\mathbf{P}^{(l+1)}(t_k^+))$, $\mathrm{P}(\overline{\mathbf{Z}}(t_k)|\mathbf{X}(t_k))=\mathrm{N}(\overline{\mathbf{Z}}(t_k);\overline{h}(\mathbf{X}^{(l+1)}(t_k^+)),\mathbf{R}^{(l+1)}(t_k))$ and

$$\begin{aligned} q^{(l+1)}(\mathbf{X}(t_k)|\overline{\mathbf{Z}}_{1k})&=\frac{\mathrm{P}^{(l+1)}(\mathbf{X}(t_k)|\overline{\mathbf{Z}}_{1k})\mathrm{P}(\overline{\mathbf{Z}}(t_k)|\mathbf{X}(t_k))}{\int \mathrm{P}^{(l+1)}(\mathbf{X}(t_k)|\overline{\mathbf{Z}}_{1k})\mathrm{P}(\overline{\mathbf{Z}}(t_k)|\mathbf{X}(t_k))d\mathbf{X}(t_k)} \\ &\triangleq \mathrm{N}(\mathbf{X}(t_k);\mathbf{X}^{(l+1)}(t_k),\mathbf{P}^{(l+1)}(t_k)) \end{aligned}$$

The predicted cubature points are calculated:

$$\chi_i^{+(l+1)}=\mathbf{X}^{(l+1)}(t_k^+)+\sqrt{\mathbf{P}^{(l+1)}(t_k^+)}\delta_i \quad i=1,2,\dots,2n \tag{15}$$

The predicted measurement vector is calculated in (16).

$$\mathbf{Z}^{(l+1)}(t_k^+)=\frac{1}{2n}\sum_{i=1}^{2n}h(\chi_i^{+(l+1)}) \tag{16}$$

$\mathbf{P}_{\mathbf{ZZ}}^{(l+1)}$ and $\mathbf{P}_{\mathbf{XZ}}^{(l+1)}$ are calculated in (17)-(19).

$$\mathbb{Z} = \left[\frac{1}{2n}\sum_{i=1}^{2n}[h(\chi_i^{*(l+1)}) - Z^{(l+1)}(t_k^+)] \quad \mathbf{H}_k\mathbf{X}^{(l)}(t_k^+) - \mathbf{L}(t_k) \right] \quad (17)$$

$$\mathbf{P}_{\mathbf{ZZ}}^{(l+1)} = \mathbb{Z}\mathbb{Z}^T + \mathbf{R}^{(l+1)}(t_k) \quad (18)$$

$$\mathbf{P}_{\mathbf{XZ}}^{(l+1)} = \left[\mathbf{X}^{(l)}(t_k^+) - \mathbf{X}(t_k^+) \right]\mathbb{Z}^T \quad (19)$$

The Kalman gain, state vector and covariance matrix are updated in (20)-(22).

$$\mathbf{K}^{(l+1)} = \mathbf{P}_{\mathbf{XZ}}^{(l+1)} / \mathbf{P}_{\mathbf{ZZ}}^{(l+1)} \quad (20)$$

$$\mathbf{X}^{(l+1)}(t_k) = \mathbf{X}(t_k^+) + \mathbf{K}^{(l+1)}(t_k)$$
$$\times \{\bar{\mathbf{Z}}(t_k) - [\mathbf{Z}^{(l+1)}(t_k^+); \mathbf{H}_k\mathbf{X}^{(l)}(t_k^+)]\} \quad (21)$$

$$\mathbf{P}^{(l+1)}(t_k) = \mathbf{P}^{(l+1)}(t_k^+) - \mathbf{K}^{(l+1)}\mathbf{P}_{\mathbf{ZZ}}^{(l+1)}(\mathbf{K}^{(l+1)})^T \quad (22)$$

The flowchart of the VBSE is given in Table I.

TABLE I.   THE FLOWCHART OF THE PROPOSED ALGORITHM

Input: $\mathbf{X}(0), \mathbf{R}(0), \mathbf{P}(0)$
Time update: for $t = t_{k-1}: t_k$ $\quad \dot{\mathbf{X}}(t) = \mathbf{F}(t)\mathbf{X}(t) + \dot{\omega}(t)$ $\quad \dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t)\mathbf{F}^T(t) + \mathbf{Q}(t)$ end
Initialization: $\mathbf{X}^{(0)}(t_k) = \mathbf{X}(t_k^+)$, $\mathbf{P}^{(0)}(t_k) = \mathbf{P}(t_k^+)$.

Variational Bayesian measurement update:
 for $l = 0: N$
    Calculate $\mathbf{A}^{(l)}(t_k)$ and update $q^{(l+1)}[\mathbf{P}(t_k^+)]$ by using Eqs. (8) and (10a).
    Using the cubature points to update $\mathbf{B}^{(l)}(t_k)$ and $q^{(l+1)}[\mathbf{R}(t_k)]$ by using Eqs. (9)-(10b).
    Update $\mathbf{P}^{(l+1)}(t_k^+)$ and $\mathbf{R}^{(l+1)}(t_k)$ at l+1th step.
    $\mathbf{P}^{(l+1)}(t_k^+) = \{E^{(l+1)}[\mathbf{P}^{-1}(t_k^+)]\}^{-1}$
    $\mathbf{R}^{(l+1)}(t_k) = \{E^{(l+1)}[\mathbf{R}^{-1}(t_k)]\}^{-1}$
    Using Eqs. (15)-(22) to update the $\mathbf{X}(t_k)$ and $\mathbf{P}(t_k)$ at the l+1th step.
 end

Output: $\mathbf{X}^{(N)}(t_k), \mathbf{P}^{(N)}(t_k)$

The main contributions of this work are shown:

(1) The position information is augmented in the measurement vector in order to acquire the more accurate position information for AUV in the discrete time.

(2) The VBSE use the multi-model method based partitioned matrix to improve the algebraic precision efficiently.

III. SIMULATIONS AND RESULTS

The normalized process and measurement error covariance matrices are defined as $\bar{\mathbf{Q}}(t)$ and $\bar{\mathbf{R}}(t_k)$. The interval [0, 1] is divided into three parts approximately. Where [0, a] denotes the small effect on noise, [a, b] denotes the medium effect on noise, and [b, 1] denotes the large effect on noise. The true noise matrices are defined as follows:

$$\mathbf{Q}(t) = \begin{cases} \bar{\mathbf{Q}}(t) & \mathbb{P}_s(0 \le k_1 < a) \\ 10\bar{\mathbf{Q}}(t) & \mathbb{P}_m(a \le k_1 < b) \\ 20\bar{\mathbf{Q}}(t) & \mathbb{P}_l(b \le k_1 \le 1) \end{cases} \quad (23)$$

where $\mathbb{P}_s(\bullet)$, $\mathbb{P}_m(\bullet)$ and $\mathbb{P}_l(\bullet)$ represent the different probabilities of bringing the small, medium and large influences to the AUV, respectively. k_1 and k_2 are set as two uncorrelated random numbers. The measurement noise matrix at the discrete time k is also defined. The interval [0, 1] is divided into three parts approximately, where [0, c] denotes the small effect on noise, [c, d] denotes the medium effect on noise, and [d, 1] denotes the large effect on noise. Where a, b, c, d are the parameters which is set in the interval.

$$\mathbf{R}(t_k) = \begin{cases} \bar{\mathbf{R}}(t_k) & \mathbb{P}_s(0 \le k_2 < c) \\ 10\bar{\mathbf{R}}(t_k) & \mathbb{P}_m(c \le k_2 < d) \\ 20\bar{\mathbf{R}}(t_k) & \mathbb{P}_l(d \le k_2 \le 1) \end{cases} \quad (24)$$

The simulation results from time 1 to 120 with a process time step of 0.01s and observation sampling time of 1s, resulting in 120 sampling points. Choosing $\bar{\mathbf{Q}}(t)$ and $\bar{\mathbf{R}}(t_k)$ as (25). The position of the measurement point is set as $(\mathbb{X}_0, \mathbb{Y}_0) = (100, 200)$.

$$\begin{cases} \bar{\mathbf{Q}}(t) = 10^{-5} \times diag[1(m)^2, 0.01(m/s)^2, 1(m)^2 0.01(m/s)^2] \\ \bar{\mathbf{R}}(t_k) = 10^{-5} \times diag[1(m)^2, 1(deg)^2, 0.01(m)^2, 0.01(m)^2] \end{cases} \quad (25)$$

From Fig. 2, it is easy to find that the VBSE makes better performance than other three traditional algorithms. Moreover, to evaluate the different algorithms more accurately, the root mean squared error (RMSE) and average RMSE (ARMSE) are defined in (26)-(27):

$$\text{RMSE}(t_k) = \sqrt{(X(t_k) - X_T(t_k))^2 + (Y(t_k) - Y_T(t_k))^2} \quad (26)$$

$$\text{ARMSE} = \frac{1}{n}\sum_{k=1}^{n}\sqrt{(X(t_k) - X_T(t_k))^2 + (Y(t_k) - Y_T(t_k))^2} \quad (27)$$

where $X_T(t_k)$, $Y_T(t_k)$ represents the true position of AUV at the discrete time k in x and y directions, respectively.

To evaluate the robustness of the proposed VBSE algorithm, the Monte Carlo (MC) approach is employed to compare the proposed algorithm with other traditional algorithms. The MC ARMSE is defined as follows:

$$\text{MC ARMSE} = \frac{1}{m}\sum_{j=1}^{m}\text{ARMSE} \quad (28)$$

where the number of simulations m is set equal to 20.
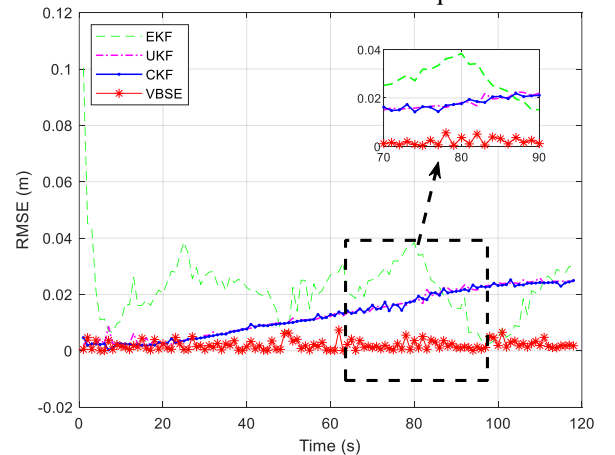
Fig.2. The RMSE for different algorithms

From TABLE II and Fig. 3, it can be seen that the proposed VBSE has the lowest MC ARMSE, this phenomenon not only represents the VBSE has higher accuracy than other traditional algorithms, but also denotes that the VBSE has a good robustness. To further compare the

performance among those four algorithms, it can be seen form TABLE II that VBSE reduces the MC ARMSE by about 91.60%, 85.15% and 85.02% compared with EKF, UKF and CKF, respectively.
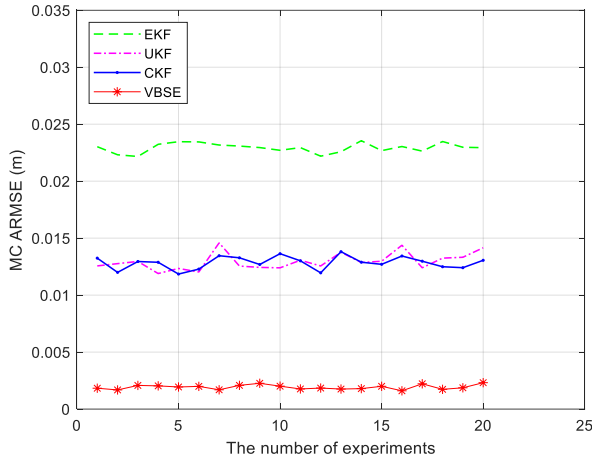
Fig. 3. MC ARMSE for different algorithms

TABLE II.   THE MC ARMSE IN MONTE CARLO EXPERIMENTS

Algorithms	EKF	UKF	CKF	VBSE
ARMSE ($10^{-3}m$)	22.90	13.00	12.90	1.93

IV.   CONCLUSIONS

The harsh marine environment inevitably makes the accurate positioning hard for MEMS grade navigation system carried in AUV. Moreover, the sample frequency of measurement update is one percent of the sample frequency of time update, so that it is not conductive to acquire the position of AUV with high accuracy. It is difficult to obtain the accurate state estimation for the traditional algorithms. The performance of the VBSE is evaluated from the perspectives of theory and simulations. It is drawn that the VBSE can determine the position of the AUV more accurately than the other three comparison algorithms, even if the measurement errors are changeable. The estimation and the error covariance matrix of states can be gained with higher accuracy by the VBSE. Moreover, not only are the accuracy and efficiency achieved by the VBSE, but also the VBSE has better robustness in the aspect of position estimation.

REFERENCES

[1] D. L. Rudnick, C. C. Eriksen, D. M. Fratantoni, M. J. Perry, "Underwater Gliders for Ocean Research," Marine Technology Society Journal, vol. 38, no. 2, pp. 73-84, 2004.

[2] H. Q. Huang, X. Y. Chen and J. J. Zhang, "Weight self-adjustment Adams implicit filtering algorithm for attitude estimation applied to underwater gliders," IEEE Access, vol. 4, no. 1, pp. 5695-5709, 2016.

[3] H. Q. Huang, J. C. Tang, B. Zhang, J. F. Chen, J. J. Zhang and X. Song, "A Novel Nonlinear Algorithm for Non-Gaussian Noises and Measurement Information Loss in Underwater Navigation," IEEE Access, vol. 8, pp. 118472-118484, 2020.

[4] S. A. Jenkins and G. D'Spain, "Autonomous underwater glider," in Springer Handbook of Ocean Engineering Springer, Dordrecht/ Heidelberg/ London/ New York (2016).

[5] J. L. Crassidis , Kok-Lam Lai and R. R. Harman, "Real-Time Attitude-Independent Three-Axis Magnetometer Calibration," Journal of Guidance, Control, and Dynamics, vol. 28, no. 1, pp. 114-117, 2005.

[6] H. Q. Huang, X. Y. Chen, B. Zhang, and J. Wang, "High accuracy navigation information estimation for inertial system using the multi-model EKF fusing adams explicit formula applied to underwater gliders," ISA Transactions, vol. 66, pp. 414-424, 2017.

[7] F. Y. Liu, X. H. Sun, Y. F. Xiong, H. Q. Huang, X. T. Guo, Y. Zhang and C. Shen, "Combination of iterated cubature Kalman filter and neural networks for GPS/INS during GPS outages," Review of Scientific Instruments, vol. 90, no. 12, pp. 1-10, 2019.

[8] H. L. Cao, Y. J. Zhang, Z. Q. Han, X. L. Shao, J. Y. Gao, K. Huang, Y. B. Shi, J. Tang, C. Shen, and J. Liu, "Pole-zero-temperature compensation circuit design and experiment for dual-mass MEMS gyroscope bandwidth expansion," IEEE/ASME Transactions on Mechatronics, vol. 24, no. 2, pp. 677-688, 2019.

[9] Z. Ghahramani, Learning Dynamic Bayesian Network. Adaptive Processing of Temporal Information, Lecture Notes in Artificial Intelligence. Springer-Verlag (1997).

[10] Y. Huang, Y. Zhang, Y. Zhao, and J. Chambers, "A novel robust Gaussian-Students t mixture distribution based Kalman filter," IEEE Trans. Signal Process, vol. 67, no. 13, pp. 3606-3620, 2019.

A Novel INS/CNS/GNSS Integrated Navigation Algorithm

Di Liu
School of Instrument Science and Engineering
Southeast University
Nanjing, China
sdili_liudi@163.com

Xiyuan Chen*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
chxiyuan@seu.edu.cn

Xiao Liu
School of Instrument Science and Engineering
Southeast University
Nanjing, China
sidescan@126.com

Zhe Yan
School of Instrument Science and Engineering
Southeast University
Nanjing, China
seuyanzhe@163.com

Abstract—**For improving the navigation ability of the INS/CNS/GNSS system in the non-Gaussian noise environment, a novel navigation algorithm based on maximum correntropy generalized high-degree CKF (MCHCKF) is proposed. First, the INS/GNSS and INS/CNS systems obtain the sub-states by the MCHCKF algorithm, respectively. Then, according to the minimum variance criterion and cubature rule, the sub-state estimations are fused to get the global state. Finally, the superior performance of the proposed method is proved in the missile-borne navigation simulation system.**

Keywords—INS/CNS/GNSS, high-degree CKF, navigation algorithm

I. INTRODUCTION

INS/CNS/GNSS system adopts Kalman filtering technology for data fusion mainly in two methods: the first is the centralized Kalman filter, the second is the decentralized Kalman filter [1]. The first filter can obtain optimal state estimation, but it has poor real-time and fault tolerance [2,3]. The second filter has small computation and good fault tolerance, so it has received extensive attention [4]. Meng et al. [5] proposed a UKF-based fusion method for INS/CNS/GNSS system. This algorithm overcomes the shortcomings of the federated Kalman filter and its improved algorithm and can obtain the global optimal state estimation. The above traditional centralized Kalman filter and decentralized Kalman filter have good performance under Gaussian assumption. However, the INS/CNS/GNSS system is easily influenced by the complex external environment, and the system measurement noise is non-Gaussian [6]. In this case, the performance of the centralized Kalman filter and decentralized Kalman filter will be seriously degraded [7,8].

The contribution of our article include the following: (1) To get the global optimal state for the INS/CNS/GNSS system under non-Gaussian measurement noise, and by following our previous work [9], a novel navigation algorithm based on MCHCKF is proposed; (2) The algorithm uses the MCHCKF to suppress the non-Gaussian noise so that the sub-navigation system obtains the optimal sub-state estimations; (3) According to the minimum variance criterion and cubature rule, the sub-state are fused to obtain the optimal global state.

The structure of the rest is: INS/CNS/GNSS system model is introduced in Section II; a novel INS/CNS/GNSS navigation algorithm is proposed in Section III; In Section IV, experiments and analysis are presented. Finally, the conclusion is given in Section V.

II. INS/CNS/GNSS SYSTEM
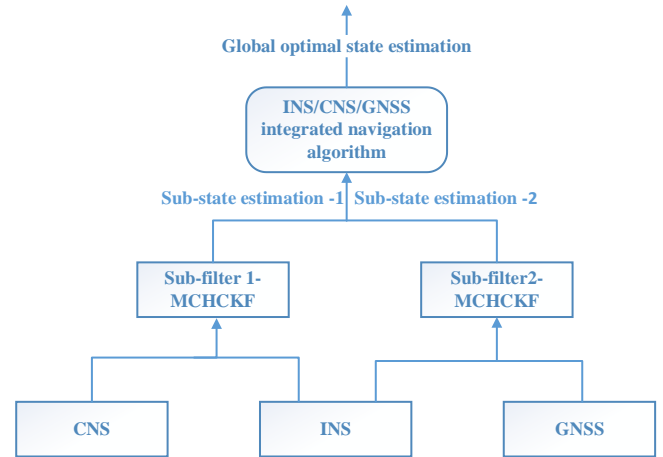
Fig. 1. The model of INS/CNS/GNSS system.

The model of the INS/CNS/GNSS system is shown in Fig. 1. In this system, we adopt the error equations of INS as the state equation, the difference between the high precision velocity and position information provided by GNSS and the velocity and position information provided by INS is used as the measurement information of the INS/GNSS sub-system, the difference between the attitude output by CNS and the attitude output by INS is used as the measurement information of the INS/CNS subsystem. Finally, the proposed navigation algorithm is used to fuse the sub-state to compensate for the system errors of INS.

Taking the launch coordinate system as the navigation system, and we can get the following state equation [9].

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1} \tag{1}$$

In the INS/CNS sub-system, the difference of attitude angle between INS and CNS output is used as measurement, and the equation of measurement of the sub-system is,

$$\mathbf{z}_{1,k} = \mathbf{H}_1 \mathbf{x}_k + \boldsymbol{\omega}_1 \tag{2}$$

---

*Xiyuan Chen is the corresponding author. (e-mail: chxiyuan@seu.edu.cn).

where, $\mathbf{z}_{1,k}$ represents the measurement vector; $\mathbf{H}_1$ denotes the measurement matrix and $\mathbf{H}_1 = [\mathbf{I}_{3\times3}, \mathbf{0}_{3\times12}]$; $\boldsymbol{\omega}_1 = [r_1, r_2, r_3]^T$ is the measurement noise and its variance is $E(\boldsymbol{\omega}_1 \boldsymbol{\omega}_1^T) = \mathbf{R}_1$.

In INS/GNSS sub-system, the differences of velocity and position between INS and GNSS output are served as measurements, and the equation of measurements of the sub-system is,

$$\mathbf{z}_{2,k} = \mathbf{H}_2 \mathbf{x}_k + \boldsymbol{\omega}_2 \qquad (3)$$

where, $\mathbf{z}_{2,k}$ represents the measurement vector; $\mathbf{H}_2$ denotes the measurement matrix and $\mathbf{H}_2 = [\mathbf{0}_{6\times3}, \mathbf{I}_{6\times6}, \mathbf{0}_{6\times6}]$; $\boldsymbol{\omega}_2 = [r_4, r_5, r_6, r_7, r_8, r_9]^T$ is the measurement noise and its variance is $E(\boldsymbol{\omega}_2 \boldsymbol{\omega}_2^T) = \mathbf{R}_2$.

III. INTEGRATED NAVIGATION ALGORITHM BASED ON MCHCKF

The proposed navigation method for the INS/CNS/GNSS system in this paper is divided into sub-state estimation and global state optimal estimation. In the sub-state estimation process, we use the MCHCKF to constrain the non-Gaussian noise and obtain the optimal local state. For the i-th $(i = 1, 2)$ filter, the specific sub-state estimation process is described in our previous study [9].

After each subsystem complete the filtering based on MCHCKF, we can obtain the sub-state estimations $\mathbf{x}_{i,k|k}$ and corresponding covariance matrix $\mathbf{P}_{i,k|k}$. Then, we obtain the global optimal state based on the principle of minimum variance and cubature rule:

$$\hat{\mathbf{x}}^*_{k|k} = \sum_{i=1}^2 \boldsymbol{\beta}_i \hat{\mathbf{x}}_{i,k|k} \qquad (4)$$

Here, $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]^T$, $\boldsymbol{\beta} = \Sigma_{k|k}^{-1} \cdot \bar{\mathbf{E}} \cdot (\bar{\mathbf{E}}^T \cdot \Sigma_{k|k}^{-1} \cdot \bar{\mathbf{E}})^{-1}$, $\bar{\mathbf{E}} = [\mathbf{I} \quad \mathbf{I}]^T$, $\mathbf{I}$ represents the n-dimensional identity matrix, $\sum_{k|k} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}$, $\mathbf{P}_{11}$ is the state covariance obtained by the INS/CNS sub-system performing the local state estimation, $\mathbf{P}_{22}$ is the state covariance obtained by the INS/GNSS sub-system performing the local state estimation, $\mathbf{P}_{12}$ and $\mathbf{P}_{21}$ denote the covariance matrix between the estimated states of the subsystems, and they are approximated by the cubature criterion:

$$\mathbf{P}_{12} = \sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{X}_{1,k|k-1}^{j*} - \mathbf{x}_{1,k|k-1})(\mathbf{X}_{2,k|k-1}^{j*} - \mathbf{x}_{2,k|k-1})^T$$
$$- \left(\sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{X}_{1,k|k-1}^{j*} - \mathbf{x}_{1,k|k-1})(\mathbf{H}_2 \mathbf{X}_{2,k|k-1}^{j} - \hat{\mathbf{z}}_{2,k|k-1})^T \right) \mathbf{K}_{2,k}^T$$
$$- \mathbf{K}_{1,k} \left(\sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{H}_1 \mathbf{X}_{1,k|k-1}^{j} - \hat{\mathbf{z}}_{1,k|k-1})(\mathbf{X}_{2,k|k-1}^{j*} - \mathbf{x}_{2,k|k-1})^T \right)$$
$$+ \mathbf{K}_{1,k} \left(\sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{H}_1 \mathbf{X}_{1,k|k-1}^{j} - \hat{\mathbf{z}}_{1,k|k-1})(\mathbf{H}_2 \mathbf{X}_{2,k|k-1}^{j} - \hat{\mathbf{z}}_{2,k|k-1})^T \right) \widehat{\mathbf{K}_{2,k}^T} \qquad (5)$$

$$\mathbf{P}_{21} = \sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{X}_{2,k|k-1}^{j*} - \hat{\mathbf{x}}_{2,k|k-1})(\mathbf{X}_{1,k|k-1}^{j*} - \hat{\mathbf{x}}_{1,k|k-1})^T$$
$$- \left(\sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{X}_{2,k|k-1}^{j*} - \hat{\mathbf{x}}_{2,k|k-1})(\mathbf{H}_1 \mathbf{X}_{1,k|k-1}^{j} - \hat{\mathbf{z}}_{1,k|k-1})^T \right) \mathbf{K}_{1,k}^T$$
$$- \mathbf{K}_{2,k} \left(\sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{H}_2 \mathbf{X}_{2,k|k-1}^{j} - \hat{\mathbf{z}}_{2,k|k-1})(\mathbf{X}_{1,k|k-1}^{j*} - \hat{\mathbf{x}}_{1,k|k-1})^T \right)$$
$$+ \mathbf{K}_{2,k} \left(\sum_{j=1}^{2n^2+1} \mathbf{w}_j (\mathbf{H}_2 \mathbf{X}_{2,k|k-1}^{j} - \hat{\mathbf{z}}_{2,k|k-1})(\mathbf{H}_1 \mathbf{X}_{1,k|k-1}^{j} - \hat{\mathbf{z}}_{1,k|k-1})^T \right) \widehat{\mathbf{K}_{1,k}^T} \qquad (6)$$

IV. SIMULATION EXPERIMENT AND RESULT ANALYSIS

The algorithm we proposed is used in the missile-borne INS/CNS/GNSS simulation system and compared with the centralized filter algorithm and the pure INS algorithm.

Firstly, we simulate the trajectory of the missile, which is shown in Fig. 2. The settings of the trajectory parameters follow the recommendations of [9], and the INS/CNS/GNSS system starts to work outside the atmosphere. The attitude error of the star sensor is $3''$. The horizontal position error, the height error and the velocity error of the GNSS are 1m, 3m, and 0.05m/s, respectively. The constant noise of the gyros and accelerometer is $1°/h$ and $100ug$, respectively. The random noise of the gyros and accelerometer is $0.5°/h$ and $50ug$, respectively.
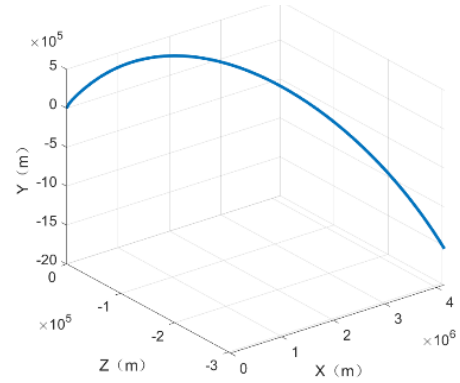
Fig. 2. The trajectory of the missile.

In the simulation, the initial state estimation, the corresponding state covariance, and the system noise variance are set to: $\mathbf{x}_{1,0|0} = \mathbf{x}_{2,0|0} = \mathbf{0}$, $\mathbf{P}_{1,0|0} = \mathbf{P}_{2,0|0} = \mathbf{diag}[(6')^2, (10'')^2, (10'')^2, (1\times10^{-1} m/s)^2 \mathbf{I}_{1\times3},$ $(5m)^2 \mathbf{I}_{1\times3}, (1°/h)^2 \mathbf{I}_{1\times3}, (1\times10^2 \mu g)^2 \mathbf{I}_{1\times3}]$, $\mathbf{Q} = \mathbf{diag}[((5\times10^{-1})°/h)^2, ((5\times10^{-1})°/h)^2,$ $((5\times10^{-1})°/h)^2, (50\mu g)^2, (50\mu g)^2, (50\mu g)^2]$. The measurement noise of each sub-system is non-Gaussian noise, which satisfies the following mixed-Gaussian noise distribution:

Measurement noise of the INS/CNS sub-system:

$$r_1 \sim 9\times10^{-1} \cdot N(0, (3'')^2) + 10^{-1} \cdot N(0, (30'')^2),$$
$$r_2 \sim 9\times10^{-1} \cdot N(0, (3'')^2) + 10^{-1} \cdot N(0, (30'')^2),$$
$$r_3 \sim 9\times10^{-1} \cdot N(0, (3'')^2) + 10^{-1} \cdot N(0, (30'')^2).$$

Measurement noise of the INS/GNSS sub-system:

$$r_4 \sim 9 \times 10^{-1} \cdot N(0,(5 \times 10^{-2} m/s)^2) + 10^{-1} \cdot N(0,(5 \times 10^{-1} m/s)^2),$$

$$r_5 \sim 9 \times 10^{-1} \cdot N(0,(5 \times 10^{-2} m/s)^2) + 10^{-1} \cdot N(0,(5 \times 10^{-1} m/s)^2),$$

$$r_6 \sim 9 \times 10^{-1} \cdot N(0,(5 \times 10^{-2} m/s)^2) + 10^{-1} \cdot N(0,(5 \times 10^{-1} m/s)^2),$$

$$r_7 \sim 9 \times 10^{-1} \cdot N(0,(1m)^2) + 10^{-1} \cdot N(0,(10m)^2),$$

$$r_8 \sim 9 \times 10^{-1} \cdot N(0,(3m)^2) + 10^{-1} \cdot N(0,(30m)^2),$$

$$r_9 \sim 9 \times 10^{-1} \cdot N(0,(1m)^2) + 10^{-1} \cdot N(0,(30m)^2).$$

Figs. 3-5 show the navigation information error curves of the missile solved by the pure inertial system (INS), the proposed algorithm, and the centralized filtering algorithm based on CKF (CCKF), UKF (CUKF), EKF (CEKF). From the figure, we can see that the navigation information error of the INS shows a decentralized trend, and its accuracy is the worst compared with the other algorithms. This is mainly because of the errors of the INS accumulates over time. CCKF, CUKF, and CEKF algorithms overcome the shortcomings of the INS, and the attitude, velocity, and position errors of their solution show a convergent trend. However, since they are suboptimal filters based on the Gaussian assumption, their performance can't achieve the desired effect under non-Gaussian noise conditions. Compared with other algorithms, our algorithm has the best performance. The main reason is that the MCHCKF algorithm has better robustness to non-Gaussian noise in the subsystem, and our algorithm can get the optimal global state according to the principle of minimum variance and cubature rule.

(c) Roll error

Fig. 3. Attitude error.

(a) Velocity-x

(a) Yaw error

(b) Velocity-y

(b) Pitch error

(c) Velocity-z

Fig. 4. Velocity error.

139

(a) Position-x

(b) Position-y

(c) Position-z

Fig. 5. Position error.

The RMSEs of the navigation information gotten by the five algorithms in the 40s-1110s is shown in Tabs. I-III. From the table, we can see that the ability of our method is better than other algorithms in terms of attitude, velocity, and position accuracy, which also shows that our algorithm has good robustness to non-Gaussian noise.

V. CONCLUSION

To improve the performance of the INS/CNS/GNSS system in non-Gaussian noise conditions, a novel navigation algorithm based on the MCHCKF is presented. Firstly, we use the MCHCKF algorithm in the subsystem to suppress non-Gaussian noise to get high-precision local state. Then, to get the global optimal state, we fuse the local state estimation according to the minimum variance and the cubature rule. Finally, the simulation experiment proves that our algorithm has good robustness under non-Gaussian noise conditions.

REFERENCES

[1] B. B. Gao, G. G. Hu, et al., "Multi-sensor optimal data fusion based on the adaptive fading unscented kalman filter," Sensors, vol. 18, no. 2, pp. 488, February 2018.

[2] Y. Gao, E. J. Krakiwsky, et al., "Comparison and analysis of centralized, decentralized, and federated filters," Navigation, vol. 40, no. 1, pp. 69-86, 1993.

[3] J. Ma, S. L. Sun, "Centralized fusion estimators for multisensor systems with random sensor delays, multiple packet dropouts and uncertain observations," IEEE Sens. J., vol. 13, no. 4, pp. 1228-1235, April 2012.

[4] J. X. Feng, Z. D. Wang, et al., "Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises," Information Fusion, vol. 14, no. 1, pp. 78-86, January 2013.

[5] Y. Meng, S. S. Gao, et al., "UKF-based optimal data fusion method for integrated INS/GNSS/CNS," Journal of Chinese Inertial Technology, vol. 24, no. 6, pp. 746-751, June 2016.

[6] B. W. Hou, Z. M. He, et al., "SINS/CNS Integrated Navigation System for Ballistic Missile based on Maximum Correntropy Kalman Filter," in Proc. IEEE American Control Conference, USA, 2018, pp. 1473-1478.

[7] B. D. Chen, X. Liu, et al., "Maximum correntropy Kalman filter," Automatica, vol. 76, pp. 70-77, February 2017.

[8] X. Liu, H. Qu, et al., "Maximum correntropy square-root cubature Kalman filter with application to SINS/GPS integrated systems," ISA Trans., vol. 80, pp. 195-202, September 2018.

[9] D. Liu, X. Y. Chen, et al., "Maximum correntropy generalized high-degree cubature Kalman filter with application to the attitude determination system of missile," Aerosp. Sci. Technol., vol. 95, pp. 105441, December 2019

TABLE I. RMSEs Of Attitude (″)

	INS	CCKF	CUKF	Proposed	CEKF
Yaw angle	834.42	4.6531	4.6537	1.3355	4.6538
Pitch angle	669.77	1.4088	1.4112	1.1067	1.4116
Roll angle	378.61	2.5029	2.5019	1.2966	2.5017

TABLE II. RMSEs Of Velocity (*m / s*)

	INS	CCKF	CUKF	Proposed	CEKF
x-direction	0.97	0.29393	0.29391	0.00831	0.29391
y-direction	3.03	0.63482	0.63481	0.01012	0.63481
z-direction	6.75	1.48029	1.48027	0.08402	1.48025

Method for Estimating the Location of A Low-frequency Target in A Shallow Sea Based on A Single Vector Hydrophone

Xianbin Sun*
School of Mechanical and Automotive Engineering
Qingdao University of Technology
Qingdao, China
Institute of Oceanographic Instrumentation
Shandong Academy of Sciences
Qingdao, China
robin_sun@qut.edu.cn

Xinming Jia
School of Mechanical and Automotive Engineering
Qingdao University of Technology
Qingdao, China
jiaxinming_123@163.com

Yi Zheng
Institute of Oceanographic Instrumentation
Shandong Academy of Sciences
Qingdao, China
850087643@qq.com

Zhen Wang
Institute of Oceanographic Instrumentation
Shandong Academy of Sciences
Qingdao, China
41434011@qq.com

Abstract—To estimate the location of a low-frequency target in the complex environment of shallow seas, both real-time properties and accuracy must be considered. In this paper, a multichannel information fusion method is proposed for estimating the location of a target based on a single vector hydrophone. First, a window signal fusion algorithm is proposed, which combines the EM algorithm to achieve adaptive signal extraction. The RNN used later realized the self-localization of the sound source. The results show that the fixed-dynamic window based on the steepest rising segment of Shannon entropy can divide a very short signal sample into a sufficient number of signal sets and the signal self-replenishment based on the EM algorithm can further strengthen the characteristics of different signal segments on the basis of replenishing the signal. Compared with other networks, the RNN has the highest accuracy and stability for this type of location estimation of an acoustic source based on the time-domain signal. The experimental results indicate that in the shallow sea environment, this model only depends on a single vector hydrophone to collect signals and rapidly locate the acoustic source, and the error radius is controlled within 1.5 m.

Keywords— shallow sea, vector hydrophone, estimating the location, EM, RNN

I. INTRODUCTION

The shallow sea environment not only has the complicated characteristics of spatiotemporal variability and uncertainty, and the signal reflected from the shallow sea floor and human activity also cause the aliasing of target signals [1]. The current method and model of acoustic field modeling are not able to correctly describe the propagation law and propagation characteristics of low-frequency signals in the shallow sea. In addition, along with the development of denoising and stealth technology, the operating frequency band of ship sonar is approaching low frequency[2]. To improve the coastal security of shallow sea areas near the continental shelf, studies on the theory and method for estimation of frequency acoustic sources (targets) in the shallow sea environment have become hot research topics in various countries.

The vector hydrophone can obtain the information of sound pressure and acoustic particle velocity in the ocean sound field at the same time and at the same point. It has the characteristics of dipole directivity and frequency independence. Therefore, a single vector hydrophone can accomplish the function that used to be realized only when multiple acoustic pressure hydrophone arrays are arranged [3]. Agarwal uses Iterative Adaptive Approach (IAA) to estimate the DOA and power of underwater acoustic emission signals using a single vector sensor, and it is found that the algorithm is robust to partial related or coherent sources in shallow sea conditions [4]. Zhao proposed an azimuth angle estimation method with a single acoustic vector sensor based on an active sonar detection system [5]. However, at present, the application of single vector hydrophone mostly needs to establish a complex mathematical model to realize DOA estimation, so it has a large dependence on the background information such as surrounding environment parameters, and cannot directly locate the location of sound source.

Because neural networks have the advantages of self-adaption, self-organization, and self-learning, they have been applied to the aspects of target identification, trend prediction, and trajectory tracking [6]. Kim and Jeong [7] determined the reference routes and navigation modes of ships through a radial basis function (RBF) network. Li, et al. [8] used the nonlinear autoregressive neural network (NARX) to predict ship movement trends based on a time sequence. In recent years, neural networks have been increasingly applied for marine ships; however, to our knowledge, previous studies have not focused on estimating the location of shallow sea targets.

As a kind of neural network, recurrent neural network (RNN) has the characteristics of connecting neurons in the same layer and is more suitable for dealing with time series problems[9]. It has been widely used in text translation[10] and speech recognition[11].

In this paper, we propose a location estimation method that uses a fixed-dynamic window to fuse the signals, apply the EM algorithm to replenish the signal and use a RNN to estimate the location. Through a single vector hydrophone,

---

*Xianbin Sun is the corresponding author. (e-mail: robin_sun@qut.edu.cn).

this method can rapidly and accurately estimate the location of a frequency acoustic source (target) in a shallow sea environment.

The model is data-driven, that is, the big data training model can replace the complex environment parameter setting, so that the model has a stronger adaptability. The real-time location of sound source in low parameter condition is realized.

II. SIGNAL FUSION OF FIXED-DYNAMIC WINDOW

A. Cross-spectrum Method

As shown in Fig 1, the *O-XYZ* coordinate system is an absolute coordinate system and the *o-xyz* coordinate system is a relative coordinate system of the vector hydrophone, where the *XOY* plane is the sea level. Due to the problem of ocean current disturbance, sea floor unevenness, and manufacturing precision, a deviation angle likely occurs between the *O-XYZ* coordinate system and the *o-xyz* coordinate system. After propagation attenuation and mixing with ambient noise, the signal emitted by the acoustic source point P is received by the vector hydrophone at point Q. The signal channels received at point Q are the sound pressure signal p, the vibration velocity v_x in the *x*-axis direction, the vibration velocity v_y in *y*-axis direction, and the vibration velocity v_z in the *z*-axis direction. θ and φ are the pitch angle and the azimuth of the acoustic source relative to the vector hydrophone, respectively, and they take the *xoy* plane and the *x* axis as $0°$, respectively. Because the acoustic source in the environment of this paper is a far-field acoustic source, the sound pressure is approximated to be of the same phase as the vibration velocity[12, 13].
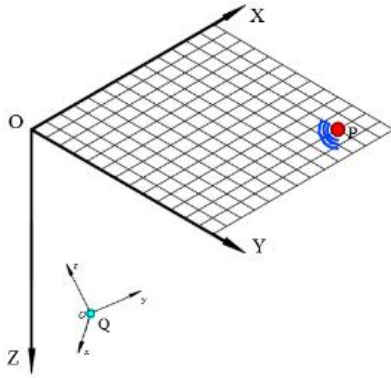
Fig. 1.   Acoustic source and the vector hydrophone

The sound pressure equation is:
$$p(\mathbf{r}, t) = A e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})} \tag{1}$$

The relationship between sound pressure and acoustic particle velocity is:
$$\mathbf{v} = -\frac{1}{\rho} \int \nabla p \, dt \tag{2}$$

In equations (1) and (2), $\mathbf{r}$ is the vector diameter; ω is the acoustic wave frequency; $\mathbf{k}$ is the wave vector; ρ is the medium density; and A is the plane wave sound pressure amplitude. Because the research situation in this paper is far-field target estimation, where $kr \gg 10$, acoustic impedance ratio $Z \approx 0$, and phase difference is small. Therefore, it can be approximated that the sound pressure and the acoustic particle velocity are in phase.

Each channel signal received by the vector hydrophone is assumed to be composed of effective signals (including non-isotropic noise) and isotropic noise, which are represented by subscripts "s" and "n", respectively.

Taking the x axis direction as an example, there is equation (3) for the sound intensity I and the acoustic particle velocity v, where the superscripted horizontal bar represents the mean value.
$$\overline{I_x} = \overline{p(t) \cdot v_x(t)}$$
$$[p_s(t) + p_n(t)] \cdot [v_{xs}(t) + v_{xn}(t)] \tag{3}$$
$$= \overline{p_s(t) \cdot v_{xs}(t)}$$

Calculate the cross-correlation function of sound pressure p and acoustic particle velocity v in the x axis direction:
$$R_{pv_x} = \int_0^T p(t)p(t - \tau) \cos\theta \cos\varphi dt \tag{4}$$

Because the signal collected by the hydrophone Q is a discrete signal, the cross-correlation function R_{pv_x} in equation (4) is subjected to a fast Fourier transform to obtain a cross-spectrum function:
$$S_{pv_x} = S_p(f) \cos\theta \cos\varphi \tag{5}$$

Similarly, the cross-spectral functions of the sound pressures p, v_y, and v_z are:
$$S_{pv_y} = S_p(f) \cos\theta \sin\varphi \tag{6}$$
$$S_{pv_z} = S_p(f) \sin\theta \tag{7}$$

Therefore, the target elevation angle θ and azimuth angle φ are:
$$\theta = \arctan\left(\frac{S_{pv_z}(f)}{\sqrt{S_{pv_x}^2(f) + S_{pv_y}^2(f)}}\right) \tag{8}$$
$$\varphi = \arctan\left(\frac{S_{pv_y}(f)}{S_{pv_x}(f)}\right) \tag{9}$$

Combining equation (3-9), the vector hydrophone sound intensity I can be calculated as:
$$\overline{I} = \left(\overline{I_x}\cos\theta\cos\varphi + \overline{I_y}\cos\theta\sin\varphi + \overline{I_z}\sin\theta\right)$$
$$= \overline{p(t) \cdot v_x(t)}\frac{\mathrm{Re}[S_{pv_x}]}{\mathrm{Re}[S_p^2(f)]} + \overline{p(t) \cdot v_y(t)}\frac{\mathrm{Re}[S_{pv_y}]}{\mathrm{Re}[S_p^2(f)]} +$$
$$\overline{p(t) \cdot v_z(t)}\frac{\mathrm{Re}[S_{pv_z}]}{\mathrm{Re}[S_p^2(f)]} \tag{10}$$

Equation (10) shows that a correlation occurs between the direction of arrival (DOA) of the sound intensity signal and that of the acoustic source signal and the sound pressure signal. When the track of the acoustic source (target) changes, its DOA and sound pressure both will change correspondingly, thus affecting the sound intensity signal. Therefore, the location of the target acoustic source can be estimated by analyzing the sound intensity signal. Meanwhile, Equation (3) indicates that a correlation occurs between the sound pressure signal and the vibration velocity signal of various directions; therefore, the multichannel sound pressure and vibration velocity signals can be fused into a single channel of sound intensity signal to achieve data compression and reduce the computation cost.

B. Fixed-dynamic Window Based on Shannon Entropy

Shannon entropy was initially proposed by Shannon, and it is a quantitative definition that describes a signal by relating it to energy. Shannon entropy not only indicates the amount of information needed to eliminate the uncertainty but also the amount of information likely contained by an unknown event. In this paper, Shannon entropy is defined

as the amount of information that is likely contained by an unknown event [14]. Its calculation formula is as follows:

$$Shannon(X) = -\sum_{i=1}^{m} p(x_i) \cdot \log p(x_i) \quad (11)$$

where x_i is the possible value of a random event X.

We introduce the concept of window overlap ratio η, which is a fixed value that means the coincidence degree between a fixed window $W_{f_{i+1}}$ and its preceding dynamic window W_{d_i}. The subscript of parameter i means that the parameter is under the i-th loop, and $i=0$ is the initial time moment. The sliding steps of the fixed-dynamic window are as follows:

1) For the four-channel signals collected by the single vector hydrophone, p, v_x, v_y, and v_z, the following quantities are specified: the length of fixed window l_f, the initial starting point of the window t_{f_0}, the window overlap ratio η, and the minimum length and maximum length of Shannon entropy l_0 and l_1, respectively.

2) We use the fixed window W_{f_i} with the starting point of the window and the window length as t_{f_i} and l_f, respectively, synchronously truncate the signals in the four channels, and fuse the signals in the window into an instantaneous sound intensity signal of equal length, W_{I_i}.

3) In the instantaneous sound intensity signal W_{I_i}, we truncate the dynamic window W_{d_i} with length of l_{s_i} at the signal origin, and we can consider that the length of signal in this dynamic window is as short as possible under the premise of containing a sufficient amount of information. The specific procedures for the truncation of the dynamic window are as follows.

    a) In the truncated instantaneous sound intensity signal W_{I_i}, we begin with the starting point of the signal, calculate the Shannon entropy for the full length of the dynamic window and constitute the Shannon entropy signal S_{I_i}.

    b) According to the derivative of S_{I_i}, S'_{I_i}, we search for the steepest rising segment of S_{I_i}, mark the length as l_{s_i}, and then proceed to step 4.

    c) If we could not find the steepest rising segment in S_{I_i}, we can consider the signal in this fixed window W_{I_i} as the invalid signal or noise signal and mark the length as l_{s_i}. We need to conduct the following two judgments:

- If S_{I_i} is small, we consider this signal as an invalid signal and take $l_{s_i} = l_0$.

- If S_{I_i} is large, we consider this signal is a noise signal and take $l_{s_i} = l_1$.

4) Return to the second step, use the signal overlap ratio η to update the starting point $t_{f_{i+1}}$ of the fixed window $W_{f_{i+1}}$, and repeat the calculation.

Fig 2 shows the flow chart for the combined sliding of the fixed window and the dynamic window. Through the combined sliding of the fixed window and the dynamic window, we can achieve the information fusion of multichannel signals and divide the very short signal into signal segments that contain sufficient information amount,

which improves the real-time performance of subsequent procedures.
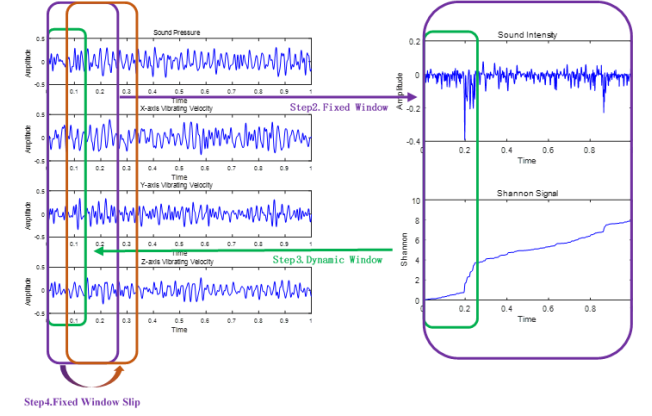
Fig. 2. Combined sliding of the fixed window and the dynamic window

C. EM Algorithm

The EM algorithm is a parameter estimation method suitable for situations with missing data. Because its every iteration is composed of an expectation step (E Step) and a maximization step (M Step), it can be described as an iterative optimization strategy [15]. The basic idea is to estimate the value of the model parameter when the initial value of missing data is given, and then estimate the value of missing data according to the parameter value. Then, according to the estimated value of missing data, we again update the parameter value and repeat the iteration until convergence.

The specific description is as follows: let Z express the missing data, i.e., not observed data, and X express the observed data, which are called incomplete data. We define the sum of the missing data Z and the incomplete data X as the complete data Y, and X is the function of Y. The following relationship is observed:

$$[L(\theta) = \ln p\,(X|\theta) \rightarrow \max] \Rightarrow [\theta \rightarrow \theta^*] \quad (12)$$

$$L(\theta) = L(x_1, \cdots, x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta) \quad (13)$$

$$\theta = \{u_i, \sigma_i^2\} \quad (14)$$

where $p(X|\theta)$ is the probability density function of the observed dataset, $p(X|\theta)$ is the probability density function of the complete dataset, and u_i and σ_i^2 are the mean and deviation of the probability density function, respectively.

The maximum likelihood function $L(\theta)$ is solved to search for θ in the parameter space Θ under the situation of fixed sampling points $\{x_1, \cdots, x_n\}$ to maximize the likelihood function:

$$\theta^* = \arg\max_{\theta \in \Theta} L(\theta) \quad (15)$$

Because $L(\theta)$ and $\ln L(\theta)$ take the extreme value at the same θ, the logarithmic likelihood function is as follows:

$$\ln\big(L(\theta)\big) = \sum_{i=1}^{n} \ln p\,(x_i; \theta) \quad (16)$$

The maximum likelihood estimation of θ, θ^*, can be solved from the following equation:

$$\frac{d}{d\theta} \ln L\,(\theta) = 0 \quad (17)$$

Therefore, Equation (9) can be written as follows:

$$L(\theta) = \sum_i \sum_{z^{(i)}} Q_i\big(z^{(i)}\big) \ln \frac{p\big(x^{(i)}, z^{(i)}; \theta\big)}{Q_i\big(z^{(i)}\big)} \quad (18)$$

$$Q_i\big(z^{(i)}\big) = p\big(z^{(i)} \big| x^{(i)}; \theta\big) \quad (19)$$

By combining Equations (18) and (19), we can obtain Equation (20) at the *t*-th iteration:

$$L\big(\theta^{(t+1)}\big) = \sum_i \ln \sum_{z^{(i)}} Q_i\big(z^{(i)}\big) \frac{p\big(x^{(i)}, z^{(i)}; \theta^{(t+1)}\big)}{Q_i\big(z^{(i)}\big)}$$

$$= \sum_i \ln \left(E\left[\frac{p\big(x^{(i)}, z^{(i)}; \theta\big)}{Q_i\big(z^{(i)}\big)}\right] \right.$$

$$\geq \sum_i \sum_{z^{(i)}} Q_i\big(z^{(i)}\big) \frac{p\big(x^{(i)}, z^{(i)}; \theta^{(t+1)}\big)}{Q_i\big(z^{(i)}\big)}$$

$$\geq \sum_i \sum_{z^{(i)}} Q_i\big(z^{(i)}\big) \frac{p\big(x^{(i)}, z^{(i)}; \theta^{(t)}\big)}{Q_i\big(z^{(i)}\big)} = L\big(\theta^{(t)}\big) (20)$$

Equation (20) can be viewed as the process to calculate the lower limit of *L(θ)*. Through the continuous iteration, we increase its lower boundary until parameter *θ* takes the maximum *θ**, and the lower boundary *L(θ(t))* converges to near the likelihood function *L(θ)* when the iteration terminates.

The specific procedures are as follows:

1) Let the number of iterations be *t=0*, and initialize the parameter vector *θ(0)*. We calculate the initial maximum likelihood function *L(0)(θ)*.

2) Step E: from *θ(t)*, we can obtain $Q_i^{(t)}\big(z^{(i)}\big)$. This step ensures that when *θ(t)* is given, the equal sign of Equation (19) is established to establish the lower boundary of *L(θ(t))*.

3) Step M: fix $Q_i^{(t)}\big(z^{(i)}\big)$, and let *θ(t)* represent the variable to calculate the derivative of *L(θ(t))* in Step E. Equation (20), we obtain *θ(t+1)*.

4) If there is $\big|L\big(\theta^{(t+1)}\big) - L\big(\theta^{(t)}\big)\big| \leq \varepsilon$, then the iteration ends, and in particular, the threshold εis the given small value. Otherwise, let $t = t + 1$ and return to Step E.

D. Recurrent Neural Network

The premise of traditional neural networks such as ANN (Artificial Neural Network) and CNN (Convolutional Neural Network) is that each element in the network, including the input and output of the network, is independent of each other. But in practical applications, because many factors are interconnected, traditional neural networks cannot handle these problems well[16, 17]. This paper uses a recurrent neural network RNN (Recurrent Neural Network) based on time series. Compared with the shortcomings of the other networks where the information flow can only achieve one-way propagation, the network adds a step of back propagation[9]. This propagation method improves the sensitivity of the RNN network to time, that is, every decision made will be affected by the residual information at the previous moment.

In Fig 3, *U*, *W*, and *V* are all the weight matrices of the RNN, and they respectively represent the connection between the input neural element x and the hidden neural element *S*, the self-recurrent connection of the hidden neural element *S*, and the connection of the hidden neural element S and the output neural element *O*. The left panel is the RNN,

and the right panel is the calculation diagram for the unfolding of the RNN.
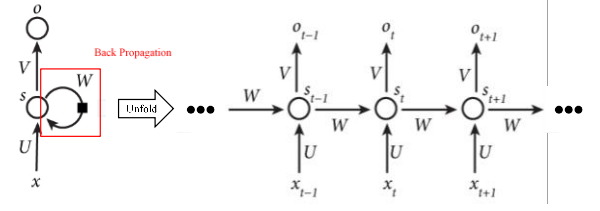
Fig. 3. Diagram for the calculation of the recurrent neural network (the input sequence of x value is projected to the corresponding sequence of the output y).

At any time *t*, the forward propagation has the following update equation:

$$s_t = f(Ux_t + Ws_{t-1} + b_t) \quad (21)$$

$$o_t = Vs_t + c \quad (22)$$

$$y_t = g(o_t) = g(Vs_t + c) \quad (23)$$

where s_t can be considered the memory at time *t*, and it is related to the memory at the previous time moment s_{t-1}; *f* is the activation function of the hidden element; *g* is the output activation function; and *b* and *c* are the offset vectors of parameters. In particular, the deviation between y_t and the actual value is the cost function *L(t)*.

In the back-propagation process of Fig 3, this network has the following gradient update equations:

$$\nabla_c L = \sum_t \left(\frac{\partial o_t}{\partial c}\right)^{\mathrm{T}} \nabla_{o_t} L \quad (24)$$

$$\nabla_b L = \sum_t \mathrm{diag}(1 - (s_t)^2)\nabla_{s_t} L \quad (25)$$

$$\nabla_V L = \sum_t \big(\nabla_{o_t} L\big)s_t^{\mathrm{T}} \quad (26)$$

$$\nabla_W L = \sum_t \mathrm{diag}(1 - (s_t)^2)\big(\nabla_{s_t} L\big)s_{t-1}^{\mathrm{T}} \quad (27)$$

$$\nabla_U L = \sum_t \mathrm{diag}(1 - (s_t)^2)\big(\nabla_{s_t} L\big)x_t^{\mathrm{T}} \quad (28)$$

We can see from Equations (20-28) that the output value of the RNN y_t is related to the previous input values $x_t, x_{t-1}, x_{t-2}, \cdots$; therefore, this network can achieve the mutual connection of output and input elements based on the time sequence through forward propagation and back propagation.

E. Overall framework of this model

In this paper, we first propose the concept of fixed-dynamic window based on the steepest rising segment of Shannon entropy. Fig 4 shows the flow chart.
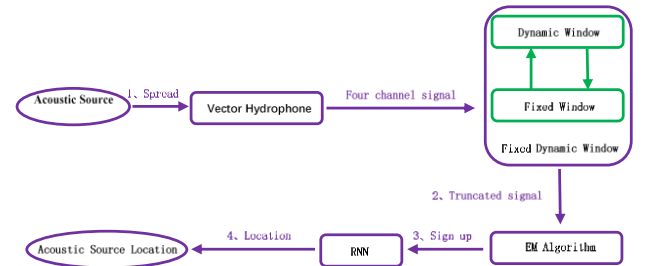
Fig. 4. Flow chart of the model.

That is, the fixed window slides synchronously in various signal channels, and the cross-spectrum method is used to obtain the instantaneous sound intensity of the target. Then, a self-adaptive dynamic window is established based on Shannon entropy to truncate the steepest rising segment of Shannon entropy in the sound intensity signal. Through this fixed-dynamic window, adequate information can be embedded in a signal that is as short as possible. The EM algorithm is then used to achieve the self-replenishment of the truncated signals, which makes the replenished signals have the same length for the ease of the subsequent calculations and amplifies the characteristic differences between various signals to improve the resolution. Finally, after training the RNN on a large number of samples, we can accurately estimate the target location for the given received signal.

III. Experiment

The data in this paper were obtained from a see-saw experiment in a shallow sea area of the South China Sea on August 2, 2018, and the acoustic source signal was a mixture of low-frequence signals at 23 Hz, 27 Hz, 33 Hz, 39 Hz, and 42 Hz. The acoustic source (target) moves at the low velocity of 4 km/h at the 5 m depth underwater, and the real-time position and track are given by a GPS signal. The single vector hydrophone was deployed at 65 m depth underwater, and the sampling frequency was 1,024 Hz.

A. Experiment of EM Algorithm

For the comparison, we use the deletion method, zero imputation, mean imputation, and KNN method to replenish the unknown data, and the signal set used by all the imputation algorithms is the same.

To quantify the imputation effect of different methods on the truncated signal, we introduce the following indexes: signal evenness (Evenness), signal similarity (Structural Similarity Index, SSIM), the average distance of various signal segments in the signal set Ω, and the useful degree of signal (Useful). In particular, the signal evenness is used to determine the noise level of the signal, the signal similarity is used to judge the similarity degree between the replenished signal and the complete signal, and the average distance of signal is an index that judges the difference of different signal segments within the signal set:

$$Evenness(X) = \sqrt{\sum_{M-1}\left(x_{(i+1)(j)} - x_{(j)}\right)^2}$$
(29)

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$
(30)

$$\Omega = \frac{\sum_i^N \sum_j^M \frac{\left(x_{(i+1)(j)} - x_{(i)(j)}\right)^2}{M}}{N}$$
(31)

where M and N are the number and signal length of signal segments contained in this signal set S_Y, respectively, x is the replenished signal, y is the true signal, u_x is the mean of x, u_y is the mean of y, $x_{(i)(j)}$ is the jth sampling point of the i-th signal segment in signal set S_Y, σ_x^2 is the variance of x, σ_y^2 is the variance of y, σ_{xy} is the covariance of x and y, and C_1 and C_2 are two constants.

Shannon entropy reflects the amount of information contained by this signal segment as well as the noise degree of this signal segment. For the convenience of

comprehensive comparison, we combine Equations (7) and (29) to propose the concept of a useful component (Useful) in the signals of the sample set.

$$Useful(S_Y) = \frac{Shannon(S_Y)}{Evenness(S_Y)} - \frac{1}{N}\sum_i^N \frac{\sum_j^M p(x_{(i)(j)}) \cdot \log p(x_{(i)(j)})}{\sqrt{\sum_{M-1}(x_{(i+1)(j)} - x_{(j)})^2}}$$
(32)

Table I shows that although the deletion method can increase the distance between individual signal segments to some extent, the information contained by the signal segment will be reduced and the useful component in the signal is significantly lost. Compared with the deletion method, the zero imputation method is worse because it reduces the distance between different signal segments and increases the difficulty in identification. By using mean imputation, all indexes are good, and it is better than the deletion method and zero imputation. For the KNN method, the distance between signals reaches the maximum (compared with all methods except for the deletion method), but the useful signal amount is significantly less than that of mean imputation. For the EM method, all indexes are good compared with the other methods except for the relatively low distance between signal segments, and it improves the useful component of signal compared with the complete signal.

TABLE I.   Results Of Different Interception Methods

Imputation method	Ω	Shannon entropy	Evenness	Useful	SSIM
Complete signal	0.42	49.47	6.16	9.18	1
Truncated signal	0.44	19.67	5.17	4.41	NA
Deletion method	0.45	19.96	5.03	4.58	NA
Zero imputation	0.19	20.78	5.42	4.29	0.02
Mean imputation	0.25	41.48	5.38	8.84	0.06
KNN	0.31	29.49	7.26	4.37	NA
EM	0.24	45.94	5.34	9.68	0.08

The goal for the replenishment of the truncated signal in this paper is to simplify the complexity degree of subsequent procedures and improve the computation speed. For the approximation degree of the replenished signal to the original signal, there is no excessively high requirement; therefore, the EM method is selected for the replenishment of the truncated signal.

B. Sound Source Position Estimation Experiment

We divide the sample set of the signal segment after performing EM imputation into a training set, a validation set, and a test set, and we use the RNN to estimate the location of the acoustic source. The network structure is the 6×1×1 three-layered network structure. The training results are Fig.5.

Fig 5 shows that the performance of RNN is satisfactory for estimating the location of the shallow-sea acoustic source. The output trajectory of the training set can fairly well fit the real motion track after many iterations, and there is only partial small deviation between the estimated trajectory of the verification set and that of the test set and the real trajectory of motion. The estimate can still be considered to have extremely high accuracy. In particular, for the azimuth estimate, the RNN can control the error to

within ±0.05°. For the distance estimate, the error is ±0.2 m. These results could be related to the ability of the vector hydrophone to estimate the azimuth of the acoustic source but inability to estimate the distance of the acoustic source.
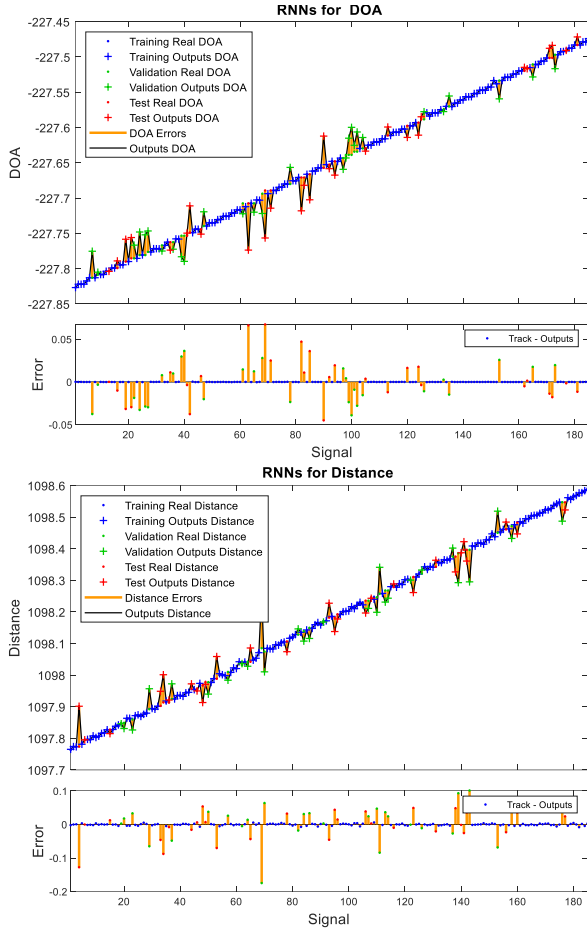
Fig. 5.   RNN training results.

To compare the estimation results of different networks on the acoustic source location, we introduce the back propagation (BP) network, RBF networks, TDNN, and NARX model to compare with the RNN adopted in this paper. The number of layers, the number of neural elements in the hidden layer, the training set, and the test set are the same for all the neural networks, and the final target for the training of different neural networks are the same. Fig 6 shows the comparison of training results for different networks, and Fig 7 shows the deviation between the test results of different networks and the real results.

By combining Figs 6 and 7, we can see that in the azimuth and distance estimates of the acoustic source (target), the performance of all networks reaches the applicable requirement. In particular, the RNN performs the best while the RBF network is the worst. The estimated location by the RBF network almost does not change with the varying trend. The performance of the TDNN follows, and its training time is longer than that of the other networks. For the NARX network, except for the shorter training time in comparison with the TDNN, the performance of various indexes is poor. Therefore, in the estimation model of acoustic source (target) location in the shallow sea, the RNN is more suitable than the other networks. After the training, it can control the accuracy of target location within 1.5 m, which could be related to the correlation between the signal

received at a certain point by the vector hydrophone and the signal within a certain time-domain radius of this point because the feedback propagation design of RNN complies more with this description.
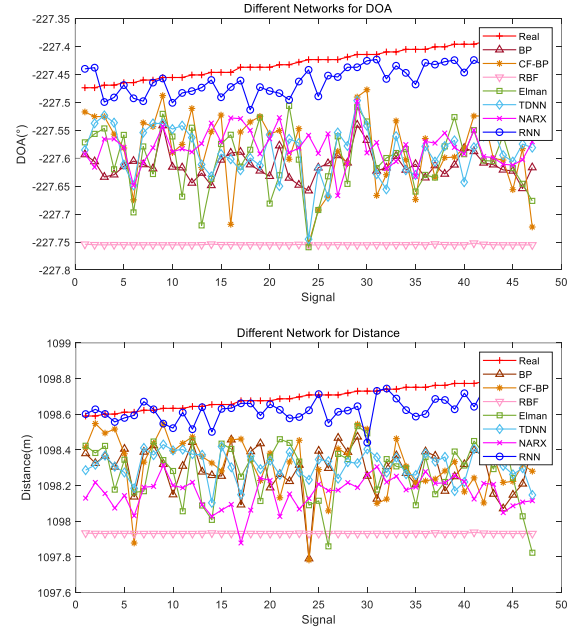
Fig. 6.   Computation results of the test set for different networks.
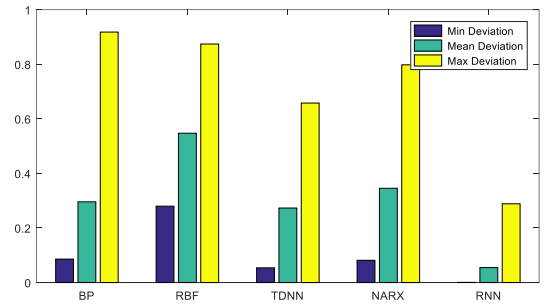
Fig. 7.   Deviation between the test results of different networks and the real results..

IV. CONCLUSIONS

In this paper, we propose a model for estimating the location of shallow-sea acoustic sources based on the truncation of signals with a fixed-dynamic window, replenishment of the signal with the EM algorithm, and training and estimation by the RNN. This model only needs to use a single vector hydrophone to collect signals from the acoustic source and achieves the fusion of four-channel signals through the fixed window, and it then uses the dynamic window to truncate the signal. Due to the problem of different lengths of truncated signals, the EM algorithm is used to achieve the self-replenishment of truncated signals. At last, for the replenished signal, the RNN is used for the training and estimation. The experimental results led to the following conclusions.

1) In comparison with the traditional acoustic source estimation model that needs to deploy a complicated vector hydrophone array to receive signals, this acoustic source estimation model only needs to adopt a single vector hydrophone to collect signals.

2) Through the fixed window, the scalar signal of sound pressure and the vector signal of vibration velocity are fused into a single-channel sound intensity signal, which to some extent eliminates the influence of posture changes of the vector hydrophone itself and environmental noise in the sea. By truncating the signal with the dynamic window, the length of the signal can be shortened as much as possible and the size of the sample set can be increased under the premise that the signal in the window contains a sufficient amount of information.

3) In comparison with the deletion method, zero imputation, mean imputation and KNN, the EM algorithm has a better comprehensive performance in the self-imputation of missing data.

4) In comparison with the BP network, RBF network, TDNN, and NARX network, the time sequence-based RNN is more suitable for this kind of dynamic acoustic source location estimation problem that is time dependent. Its predictions are more accurate, and the performance is more stable.

However, in this paper, we only estimate the location of a low-velocity acoustic source (single target) in the shallow sea environment, and for the identification and tracking of high maneuvering targets and multiple acoustic sources (multi targets), further studies need to be conducted.

REFERENCES

[1] D. Wittekind and M. Schuster, "Propeller cavitation noise and background noise in the sea," Ocean Engineering, vol. 120, pp. 116-121, 2016/07/01 2016.

[2] J. Zhou et al., "Low frequency ambient noise modeling and comparison with field measurements in the South China Sea," Applied Acoustics, vol. 148, pp. 34-39, 2019/05/01/ 2019.

[3] G. L. D'Spain, J. C. Luby, G. R. Wilson, and R. A. Gramann, "Vector sensors and vector sensor line arrays: Comments on optimal array gain and detection," *The Journal of the Acoustical Society of America,* vol. 120, no. 1, pp. 171-185, 2006/07/01 2006.

[4] A. Agarwal, A. Kumar, and M. Agrawal, "Iterative adaptive approach to DOA estimation with acoustic vector sensors," in *OCEANS 2015-Genova*, 2015, pp. 1-8: IEEE.

[5] A. Zhao, L. Ma, X. Ma, and J. Hui, "An improved azimuth angle estimation method with a single acoustic vector sensor based on an active sonar detection system," *Sensors,* vol. 17, no. 2, p. 412, 2017.

[6] A. Prieto *et al.*, "Neural networks: An overview of early research, current frameworks and new challenges," *Neurocomputing,* vol. 214, pp. 242-268, 2016.

[7] J.-S. Kim and J. S. Jeong, "Extraction of reference seaway through machine learning of ship navigational data and trajectory," *Int. J. Fuzzy Log. Intell. Syst.,* vol. 17, no. 2, pp. 82–90, 2017.

[8] G. Li, B. Kawan, H. Wang, and H. Zhang, "Neural-network-based modelling and analysis for time series prediction of ship motion," *Ship Technol. Res.,* vol. 64, no. 1, pp. 30–39, 2017.

[9] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur, "Recurrent neural network based language model," in *Eleventh annual conference of the international speech communication association*, 2010.

[10] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *arXiv: Computation and Language,* 2014.

[11] Y. Miao, M. Gowayyed, and F. Metze, "EESEN: End-to-end speech recognition using deep RNN models and WFST-based decoding," in *ieee automatic speech recognition and understanding workshop*, 2015, pp. 167-174.

[12] W. Munk, P. Worcester, and C. Wunsch, Ocean acoustic tomography. Cambridge university press, 2009.

[13] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational ocean acoustics*. Springer Science & Business Media, 2011.

[14] M. S. Hughes, "Analysis of digitized waveforms using Shannon entropy," *J. Acoust. Soc. Am.,* vol. 93, no. 2, pp. 892–906, 1993/02/01 1993.

[15] G. McLachlan and T. Krishnan, *The EM algorithm and extensions*. John Wiley & Sons, 2007.

[16] M. Van Gerven and S. Bohte, "Artificial neural networks as models of neural information processing," *Frontiers in Computational Neuroscience,* vol. 11, p. 114, 2017.

[17] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition,* vol. 77, pp. 354-377, 2018.

Development of Photoelectron Emission Yield Measurement System for Metal Materials

Yu Chen*
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenyu@xjtu.edu.cn

Yan Yang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
yangyan0415@stu.xjtu.edu.cn

Guorui Huang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
huang0923@stu.xjtu.edu.cn

Hanzhi Li
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
lhz3119304518@stu.xjtu.edu.cn

Changxi Li
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
3120304022@stu.xjtu.edu.cn

Shuang Wang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
shuang@xjtu.edu.cn

Yonghong Cheng
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
cyh@xjtu.edu.cn

Abstract—Because the spacecraft in orbit are exposed to intense solar radiation, it will be affected by factors such as cosmic rays and high-energy particles, causing its material surface to be charged and damaged. And compared to the various currents captured on the surface of the spacecraft, when the incident energy is higher than the photoelectric threshold, the damage of the emitted photocurrent to the material is more serious. Because of the significant threat of particle radiation to the charged protection and operational life of spacecraft material, this paper studied the test method of photoelectron emission yield (PEEY) and developed the corresponding system. During the test, firstly, a deuterium lamp and a vacuum ultraviolet (VUV) monochromator are used to generate light. After splitting, it uses a focusing and collimating device to modulate the light into a parallel beam. Secondly, we use a photodiode to detect the incident light flux to obtain the number of incident photons. Thirdly, the collection plate is applied to realize the useful collection of electrons emitted from the material surface. The response current is read by electrometer with high accuracy and low noise. Finally, the ratio of the number of photoelectrons to the incident photons at a single wavelength is calculated to obtain the PEEY. The result shows that the minimum detectable range of the PEEY of the gold is $10^{-2} \sim 10^{-4}$ el/ph. under VUV radiation, and the reliability of the system is proved by comparative with the literature.

Keywords—ultraviolet light, photoelectron, PEEY

I. INTRODUCTION

The orbiting spacecraft will be irradiated by the high-energy cosmic rays, solar wind particle, and plasma in the space environment, and excessive charging will cause electrostatic discharge, internal charging of the material. Its energy will exceed the threshold of photoemission. As a result, photoelectron emission and secondary electron emission occur, affecting the stable and reliable operation of the spacecraft, and even causing its failure [1]. Photoelectron emission yield (PEEY) is one of the critical factors, and it influences the electrification of the spaceship. There was a lot of researches in the field of PEEY abroad. The European Space Agency has realized photoelectron emission characteristics of a variety of space conductor materials at incident photon energy range of 5-25 eV. It includes work function, reflectivity, PEEY, and photoelectron flux of material. And also discussed methods for controlling of photoemission [2]. Tokyo Metropolitan University and JAXA jointly designed a test system for the photoelectron emission coefficient under long-pulse VUV irradiation based on a D2 discharge light source and the construction of a vacuum outer narrowband filter [3]. Dowell of Stanford Linear Accelerator Center in the United States deduced the mathematical probability model of PEEY through the three-step theory of photoemission and compared it with his experimental results, which showed a high consistency [4]. In China, Chen Yu from Xi'an Jiaotong University used an S2D2 vacuum ultraviolet light source in the range of 115 nm ~ 400 nm to measure the photoemission properties of the 200nm thick Au coating on glass and CMG-100AR solar cell cover glass. Their photoelectron emission yield curves were also given [5].

In the test of PEEY, considering the attenuation and loss of vacuum ultraviolet (VUV) during the transmission process, a deuterium lamp with high output stability and a high-resolution monochromator is used to generate and split the light source. We use a photodiode to measure the incident light flux and design a photoelectron collecting device. This research is not only helpful for analyzing the charged model of spacecraft materials, but also can improve its operational stability and life, and play a guiding role in the design process of the spacecraft structure.

II. THEORETICAL BASIS

In December 1886, Hertz, a German physicist, discovered the photoelectric effect in the experimental study of electromagnetic waves. After Thomson discovered the electron in 1897, Leonard found in his experiments that the particles produced by the photoelectric emission were electrons, called photoelectrons. However, Einstein received the Nobel Prize for his quantum theory of photoemission published in 1903. It took over 50 years before a comprehensive understanding of the photoemission process was obtained [6,7,8].

When light hits the surface of the material, there are three phenomena of reflection, transmission, and absorption. And the light absorbed by the material will produce the internal photoelectric effect and the external photoelectric effect,

---

* Yu Chen is the corresponding author. (e-mail: chenyu@xjtu.edu.cn).

photoemission belongs to the external photoelectric effect. Photoemission yield $Y(h\upsilon)$ is defined as the number of photoelectrons excited by each incident photon under a given photon energy $h\upsilon$. That is, the PEEY is the production contributed by all photoelectrons with different energies [9].

$$Y(h\upsilon) = \frac{N_{ele}}{N_{pho}} \qquad （1）$$

Among them: N_{ele} is the number of emitted electrons on the surface material; N_{pho} is the number of photons incident on the surface material. The interpretation of PEEY is based on the three-step photoemission model of Spicer [10]. It provided a beneficial description for the primary and practical applications of the photoemission process. Considering photoemission as three continuous steps: a) Light transition: electrons in a solid material absorbed the energy of the incident photon and were excited to a higher energy state. b) Transport of excited electrons to the surface: the excited electrons were scattered by electron-electrons, electron-phonons, and electron-impurities in the transport process, and electron-electron scattering is dominant. c) Electron escape through the potential surface barrier: the excited electrons migrate to the surface after being scattered; only by overcoming the potential surface barrier can they escape from the solid into the vacuum and become free electrons.

III. EXPERIMENT PROCEDURE

A. Measurement System

The Measurement system is composed of a VUV monochromatic light generation module, a continuous spectrum module, and a PEEY test module, Which includes a deuterium lamp, monochromator, an optical system, ultraviolet detector, vacuum chamber, sample stage, and displacement stage. The test device is composed of a sample chamber and an analysis chamber. Each section has independent vacuum acquisition and detection functions. The schematic diagram of the PEEY test system is shown in the Fig.1.
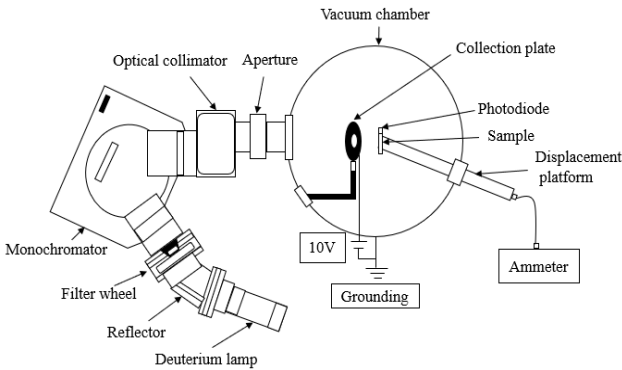
Fig. 1. PEEY measurement system.

A deuterium lamp with high stability is used as an ultraviolet light source in this paper. Its wavelength range is 115-400nm, and its brightness is six times higher than the commonly used deuterium lamp (L2D2). The light from the source is mainly focused through the reflector so that its divergence angle matches the incident angle of the monochromator. The filter is used to filter the ultraviolet light produced by second-order diffraction.

We used the monochromator for spectrophotometry. The light is projected onto the concave grating of the monochromator and exits in the form of concentrated light. As the grating rotates, monochromatic light of different wavelengths is emitted in turn, which is then collimated into parallel light by the collimator, maintaining the intensity of the emitted light. Aperture is a circular device with a hole for adjusting the size of the emission spot after the light source is collimated. Silicon photodiode is used to measure the incident light flux with an effective area of 100 mm^2 and a dark current of (typical value) 2nA.

B. Measurement Principle and Condition

The measurement principle of the PEEY is as follows. The ultraviolet light emitted by the deuterium lamp is decomposed into monochromatic light of different wavelengths by a monochromator. The aperture control diameter of the incident spot after the monochromatic light is collimated. The sample to be tested is irradiated with monochromatic light of different wavelengths. The collecting device collects the photoelectrons emitted from material to obtain the photoelectron current. At the same time, the photocurrent signal of the standard photodiode is tested under the same conditions. Then the two current signals are converted and compared to obtain the PEEY of the sample to be tested.

Before the test, the sample surface should be kept clean to simulate the material used in real space. This paper guarantees the surface flatness by spraying gold on stainless steel. And held at room temperature for some time and cleaned of surface dust with rubber air before being placed in a vacuum system. Clean the sample thoroughly with an ion gun. An electric shielding device is used to avoid the escape of photoelectrons and the interference of stray electrons.

TABLE I. CHARACTERISTIC PARAMETERS OF TEST SYSTEM

Indicator	System Parameter
Deuterium lamp	Hamamatsu L11798
Silicon photodiode	AXUV-100G
Monochromator	Model 234/302
Bias voltage	Collector 10 V; Sample 0 V
Target material	Metal material
Sample coating thickness	>100 nm Au coating
Sample size	30 mm ×30 mm× (0.1-1) mm
System vacuum	Up to 10^{-4} Pa
Three-dimensional sample stage	X, Y, Z
Sample stage drive range	XY=±10 mm, Z=100 mm
Four-dimensional test stage	X, Y, Z, R
Test stage drive range	XY=±15 mm, Z=400 mm, R=0-85°

C. Experiment Procedure

First, we use a standard photodiode to measure the number of photons. The light emitted by the deuterium lamp is split through the monochromator and irradiated on the photodiode in the vacuum chamber, and obtain current value through the ammeter. The number of incident photons np is acquired from the ratio of the light power to the incident energy.

$$P(\lambda) = \frac{I_k(\lambda)}{S_k(\lambda)} \qquad (2)$$

$P(\lambda)$ is the incident light irradiation power, $I_k(\lambda)$ is the photocurrent measured by the photodiode (minus the background noise current), and $S_k(\lambda)$ is the spectral sensitivity of the photodiode (obtained from the manufacturer's standard table).

$$E_p = h\upsilon = \frac{hc}{\lambda} \qquad (3)$$

E_p is the incident photon energy, h is Planck's constant, υ is the frequency of light waves, and c is the speed of light.

$$n_p(\lambda) = \frac{P(\lambda)}{E_p} \qquad (4)$$

n_p is the number of incident photons on the surface of the material per unit time.

The photodiode and the sample are installed on the four-dimensional displacement platform. By adjusting the coordinates of the displacement platform, the photodiode and the model are positioned at the same position in the optical path in turn. Because the photodiode is sensitive to light, the observation window on the wall of the vacuum chamber is blocked from sunlight during the test, and the vacuum gauge in the vacuum chamber is closed.
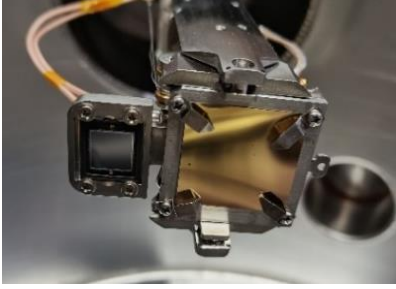
Fig. 2. Schematic diagram of photodiode and sample installation.

Secondly, the gold-plated samples were irradiated at different wavelengths in the wavelength range of 120-210 nm incident light. An electronic acquisition device is designed, and a bias voltage is applied to the device. Using the principle that electrons move in the opposite direction of the electric field, photoelectrons are collected on the collecting device. The photoelectron current is read out with a 6517B electrometer.

$$n_e(\lambda) = \frac{I(\lambda)}{e} \qquad (5)$$

n_e is the number of photoelectrons emitted per unit time on the sample surface. $I(\lambda)$ is the photoelectron.

In the vacuum chamber, a bias voltage is applied to the collector plate to realize the significant collection of photoelectrons from the surface materials. The collector plate consists of copper foil, polyethylene insulation, and a support frame. The middle hole of the collecting plate is 10 mm, placed in the center of the optical path. And close to the sample so that incident light can pass through. The copper foil and the support frame are insulated by polyethylene, and the lead is drawn by coaxial cable through a flange on the vacuum chamber. The schematic diagram is shown in the figure.
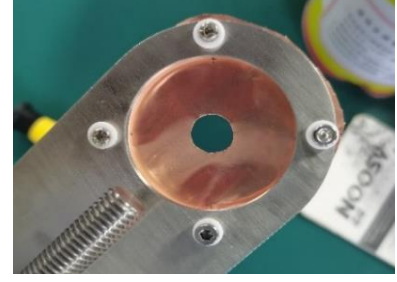
Fig. 3. Photoelectron collection device.

Finally, we obtain the PEEY by the ratio of the number of photoelectrons to the number of photons. All tests are carried out under vacuum conditions of 10^{-3} Pa and higher.

$$QE(\lambda) = \frac{n_e(\lambda)}{n_p(\lambda)} \qquad (6)$$

IV. RESULT ANALYSIS

This article tests the PEEY of the gold-plated sample on the surface of stainless steel with a deuterium lamp under the radiation wavelength of 120-210 nm. The sample size is 30 mm × 30 mm × (0.1-1) mm. A circular collection plate device is used to collect the emitted photoelectrons on the surface of the sample. We use photodiode to test the number of incident photons. The following table shows the original test data of photodiode and gold at incident wavelengths of 125 nm, 130 nm, 140 nm, 160 nm, and 210 nm.

TABLE II.　　TEST DATA

Wavelength （nm）	Photodiode current （nA）	Sample current （pA）
125	11.3	142.7
130	9.3	71.7
140	11.2	51.8
160	44	49.2
210	1.7	0.4
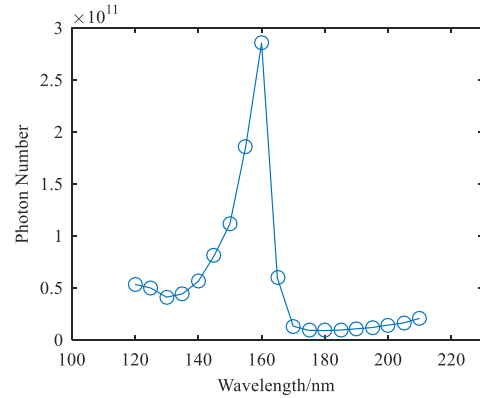
Fig. 4. The number of photons.

Fig. 4 shows the measurement results of the photodiode of the photon number in the wavelength range of incident light from 120 to 210 nm. We can see from the figure that the number of photons reaches the maximum when the wavelength of incident light is about 160 nm.
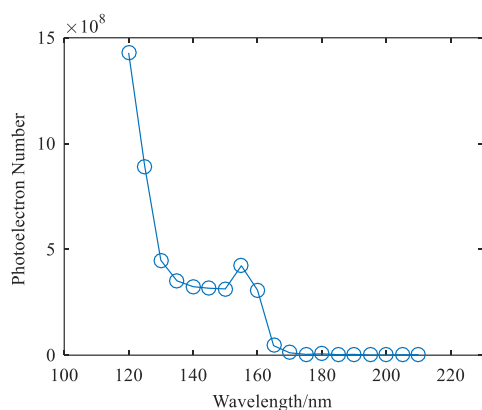
Fig. 5.   The number of photoelectrons.

Fig.5 shows the measurement results of the photoelectrons. We apply a voltage of 10 V to the collection plate to collect the photoelectrons emitted from the surface of the sample. In this paper, we obtain the photoelectron current from the sample stage. Because ideally, the current measured on the sample table and the collector plate would be the same.
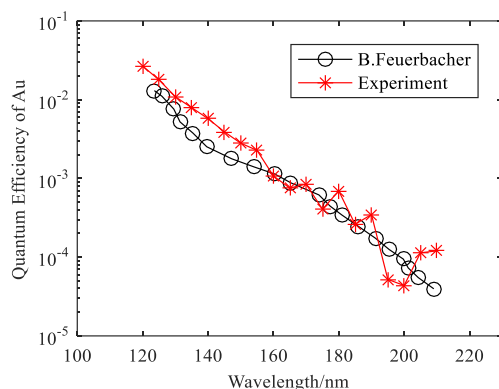
Fig. 6.   PEEY comparison between experiment and literature of gold.

In Fig. 6, the red line shows the PEEY of the gold-plated sample experimentally. The black line is the result of gold material published by B. Feuerbacher of the European Organization for Space Research in 1972 [2]. The range of the PEEY of the gold is $2.7 \times 10^{-2} \sim 1.2 \times 10^{-4}$ el/ph. By the test, B. Feuerbacher measured $1.3 \times 10^{-2} \sim 4 \times 10^{-5}$ el/ph. As can be seen from Fig. 6, the PEEY presents an overall downward trend within the test wavelength range. The reason is that with the increase of wavelength, the energy of incident light decreases, the energy absorbed on the surface of the material, and the probability of excited emission of internal electrons also decreases, so the PEEY decreases.

To analyze the two curves, B. Feuerbacher prepared the samples by placing the gold in the form of an evaporative film on glass and exposing it to the atmosphere for at least a few days. The incident photon flux was monitored with a calibrated aluminum detector near-normal incidence (7.5°), using a normal incidence grating monochromator and a DC capillary discharge. In this paper, we combine the deuterium lamp and monochromator to realize the generation and splitting of the light source, and a standard silicon photodiode to test the incident light flux. It is mounting the photodiode and the tested sample on a four-dimensional displacement platform in the vacuum chamber. Through electric control,

we place the photodiode and the model in the same position of the incident light path to complete the measurement of incident light flux and photoelectron current. The traditional measurement method of quantum efficiency adopts the double-light path method. The incident light comes out from the monochromator, and it is divided into two paths by the spectroscope. One way irradiates the photodiode, and the other way illuminates the sample to be tested. It is difficult to calculate the error introduced by the transmission and reflection of the spectroscope under different incident light. The test method in this paper ensures that the radiation flux received by the photodiode and the sample is equal, reduces the error introduced by the beam splitting, and improves the accuracy of the quantum efficiency test results and the accuracy of the test system. Although the test band in this paper is consistent with B. Feuerbacher, the test method and equipment are different, so there are some differences in the test results. Through comparison, we found that the overall trend of the two curves is consistent, and the fluctuation is small, thus proving the reliability of the test system.

V.   CONCLUSION

In this paper, we develop a PEEY test system for metal materials in the space environment by using the method that places the photoelectric detector and sample in the same plane. The system can work in a vacuum environment to collect the current signals of photodiode and gold coating samples under VUV irradiation. The PEEY of the gold coating under ultraviolet irradiation is obtained by processing and analyzing the collected signals. The test method proposed in this paper improves the accuracy of the test results and the precision of the test system. It lays a foundation for the research mechanism of metal materials under space environment irradiation.

REFERENCES

[1] K. L. Bedingfield, R. D. Leach, and M. B. Alexander, "Spacecraft system failures and anomalies attributed to the natural space environment," NASA RP-1390[R], 1996.

[2] B. Feuerbacher and B. Fitton, "Experimental Investigation of Photoemission from Satellite Surface Materials," J. Appl. Phys. vol. 43, pp. 1563-1572, 1972.

[3] K. Nitta, H. Miyake, and M. Takahashi, "Proposed secondary electron emission and photoelectron emission measurements related to engineering design optimization guidelines and spacecraft charging," 10th ISMSE, 2006.

[4] D. H. Dowell, F. K. King, and R. E. Kirby, "In-Situ Cleaning of Metal Cathodes using a Hydrogen Ion Beam," Phys. rev. St. Accel. Beams, vol. 9, pp. 257-260, 2006.

[5] Yu Chen, Atsushi Maruyama, Jiang Wu, Kazuhiro Toyoda, Mengu Cho, "Photoelectron emission yield measurement of insulator by vacuum ultraviolet light source and several narrow bandwidths filters," 12th SCTC, 2012. Kitakyushu, Japan.

[6] M. Cardona and L. Ley, "Photoemission in Solids I," Springer-Verlag, Berlin, Heidelberg, New York, 1978. G. Margaritondo, Physics Today, pp. 66, 1988.

[7] C. Coluzza, R. Sanjines, and G. Margaritondo, "Photoemission from the past to the future," Repro, Ecole Polytechnique Federale de Lousana, Switzerland, p. 1, 1992.

[8] W. E. Spicer, "Chemistry and Physics of Solid Surfaces IV," R. Vanselaw and R. Howe, Springer-Verlag, Berlin, Heidelberg, p. 1, New York, 1982.

[9] Zhantian Zhong, "Photoelectron emission yield spectrometer," Institute of Semiconductors, Chinese Academy of Sciences, 1986.

[10] W. E. Spicer, "Photoemissive, Photoconductive, and Optical Absorption Studies of Alkali-Antimony Compounds," Physical Review, vol. 112, pp. 114-122, 1958.

Analysis of Secondary Electron Yield and Energy Spectrum of Metal Materials based on Furman Model

Zecai Chen
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenzecai@stu.xjtu.edu.cn

Yu Chen*
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenyu@xjtu.edu.cn

Guorui Huang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
Huang0923@stu.xjtu.edu.cn

Changxi Li
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
3120304022@stu.xjtu.edu.cn

Qingyun Shi
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
sqy3119304516@stu.xjtu.edu.cn

Shuang Wang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
shuang@xjtu.edu.cn

Abstract—**In the study of secondary electron emission, this paper first uses the Furman model to classify secondary electrons into three types of electrons and model them separately. Second, the Monte Carlo simulation process of the model is realized. In the process of establishing Monte Carlo simulation, a single particle model is used, and the energy spectrum distribution of the entire model is presented based on the superposition of a large number of single-particle distributions. Thirdly, several sets of experiments on the corresponding energy spectrum distribution data obtained under different incident energies of one electron and different incident angles were carried out. Then, discuss the relationship between the secondary electron yield and the incident angle and energy of the primary electron, and quantitatively analyze the relationship between the secondary electron yield and the incident electron angle and incident energy. Finally, the relationship between the incident electron energy and incident angle and the secondary electron emission yield was compared and verified.**

Keywords—Secondary electron yield, Furman model, energy spectrum distribution, Monte Carlo

I. INTRODUCTION

Secondary electron emission testing technology has critical applications in many fields, such as aerospace and scanning electron microscopy. When the spacecraft is in space, it will be affected by the plasma environment of various charged particle streams. At this point, the spacecraft surface will accumulate charge and produce discharge, affecting the success of the mission.

Koons et al. studied the causes of satellite losses caused by the space environment over the past 25 years. Studies have shown that electrostatic discharge (ESD) accounts for 54.2% of all satellite accidents. Because of the above reasons, it is necessary to study the spacecraft charging caused by radioactive radiation and plasma environment, and the secondary electron emission is an essential parameter of the satellite surface charging [1].

Secondary electron emission (SEE) is usually the most critical factor affecting the current balance of charged particles in the space environment, which enables the charging of spacecraft in space plasma [2]. The spacecraft flying in orbit is affected by the flow of various charged particles in the space plasma environment. At this time, the surface of the spacecraft will accumulate charges and generate discharges. The main hazards include the short circuit of the spacecraft's power supply and distribution system and breakdown of components, thus affecting the smooth completion of the spacecraft mission. When the surface of a spacecraft in a space environment is charged, its final equilibrium potential depends on the balance between incident electrons, ions, surface photoelectron emission, backscatter current, and leakage current.

II. SECONDARY ELECTRON EMISSION MODEL

A. Basic Theory

Two important physical quantities often used in the experimental research of the secondary electron emission process are secondary electron emission yield and emission energy spectrum.

Secondary electron emission yield is the average number of secondary electrons excited by each incident electron, and it is an important parameter to measure the ability of secondary electron emission. The calculation formula of this physical quantity is as follows:

$$\delta = \frac{I_s}{I_0} \qquad (1)$$

Where, I_0 is the incident electron current, I_s is the secondary electron current, that is, the current formed by the secondary electrons emitted from the surface.

The classical secondary electron emission yield is shown in Fig.1. As the incident energy increases, δ first increases to the maximum value δ_m and then decreases. The energy of the incident electron corresponding to δ_m is denoted as E_m.
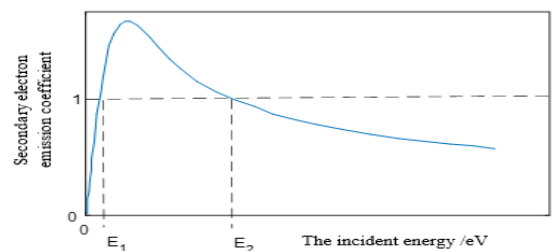
Fig. 1. Classic secondary electron emission coefficient diagram.

*Yu Chen is the corresponding author. (e-mail: chenyu@xjtu.edu.cn).

When electrons with certain incident energy bombard the sample surface, the energy distribution formed by the secondary electrons coming out of the sample surface is called the secondary electron energy spectrum distribution.

B. Furman Model

The Furman model divides the secondary electrons into three types of secondary electrons. When the stable current I_0 of electrons hits the surface, part of the elastic emission forms currently I_e, and the rest is incident into the material. Some of these electrons are scattered from one or more atoms inside the material and reflected back, these are scattered electrons, and the corresponding current is called I_r. The remaining electrons interact with the material to produce true secondary electrons, and the corresponding current is called I_{ts} [3].

1) Elastic Scattering Electron Model

The experimental data show that at vertical incidence, the following formula can be given ($\theta_0 = 0°$):

Elastic electron emission coefficient:

$$\delta_e(E_0, 0) = P_{1,e}(\infty) + \left(P_{1,e} - P_{1,e}(\infty)\right) e^{-\frac{(|E_0 - \hat{E}_e|/w)^p}{p}}$$

(1)

The probability density function of Elastic Electrons:

$$f_{1,e} = \theta(E)\theta(E_0 - E)\delta_e(E_0, \theta_0) \frac{2e^{-\frac{(E - E_0)^2}{2\sigma_e^2}}}{\sqrt{2\pi} erf\left(\frac{E_0}{\sqrt{2}\sigma_e}\right)}$$

(2)

2) Scattering Electron Model

The experimental data show that at vertical incidence, the following formula can be given ($\theta_0 = 0°$):

Emission factor:

$$\delta_r(E_0, 0) = P_{1,r}(\infty)\left(1 - e^{-\left(\frac{E_0}{E_r}\right)^r}\right)$$

(4)

The probability density function of scattered electrons:

$$f_{1,r} = \theta(E)\theta(E_0 - E)\delta_e(E_0, \theta_0) \frac{(q + 1)E^q}{E_0^{q+1}}$$

(5)

Where, $P_{1,r}(\infty)$ represents the inelastic backscattering coefficient when electrons are incident perpendicularly and the incident energy tends to infinity, E_r and r determine the changing trend of the inelastic backscattered electron emission coefficient at normal incidence, and r_1 and r_2 determine the influence of the angle on the inelastic backscattered electron emission coefficient, The role of the $\theta(x)$ function is to ensure that the value in the brackets is non-negative, thereby determining the variable range.

Satisfy the normalization condition,

$$\int_0^{E_0} dE f_{1,r}(E) = \delta_r(E_0)$$

(6)

3) True Secondary Electron Model

The energy and angle dependence of $\delta_{ts}(E_0, \theta_0)$ are well fitted experimentally by the approximate universal scaling function $D(x)$, and the expression $D(x)$ is:

$$D(x) = \frac{sx}{s - 1 + x^s}$$

(7)

For spectral functions, the following assumptions are made:

$$f_{n,ts} = \theta(E) F_n E^{p_n - 1} e^{-\frac{E}{\varepsilon_n}}$$

(8)

$$F_n^n = \frac{P_{n,ts}(E_0)}{\left(\varepsilon_n^{pn} \Gamma(p_n)\right)^n P\left(n p_n, \frac{E_0}{\varepsilon_n}\right)}$$

(9)

Where, p_n, ε_n is the phenomenon parameter, they have no actual physical meaning.

True secondary electron emission coefficient:

$$\delta_{ts}(E_0, \theta_0) = \frac{\hat{\delta}_{ts}(\theta_0)}{D\left(\frac{E_0}{\hat{E}(\theta_0)}\right)}$$

(10)

4) Probability of Launch

Even if δ_e and δ_r are constrained to satisfy $\delta_e + \delta_r \leq 1$, P_1 can exceed 1 and P_0 can become negative. For example, when $\delta_{ts} \geq 1.2$ and $\delta_e + \delta_r \geq 0.5$, a violation of the basic probability attribute may occur, which is easy to see in practice. Another definition of the emission probability of guaranteed probabilities $P_0 \geq 0$ and $P_1 \leq 1$ is to consider the probability per unit of penetrating electron current instead of per unit of incident electron current. The current that can be used to produce true secondary electrons is $I_0 - I_e - I_r$. Therefore, according to the permeation current yield is:

$$\delta_{ts}' = \frac{I_{ts}}{I_0 - I_e - I_r} = \frac{\delta_{ts}}{1 - \delta_e - \delta_r}$$

(11)

The yield of the secondary electron can be expressed by probability as follows:

$$\delta_{ts}' = \sum_{n=1}^{\infty} n P_{n,ts}'$$

(12)

The emission probability of N secondary electrons produced by one incident electron can be obtained:

$$\begin{cases} P_0 = (1 - \delta_e - \delta_r) P_{0,ts}' \\ P_1 = (1 - \delta_e - \delta_r) P_{1,ts}' + \delta_e + \delta_r \\ P_n = (1 - \delta_e - \delta_r) P_{n,ts}', n \geq 2 \end{cases}$$

(13)

It can be defined that $P_{n,ts}'$ satisfies Poisson distribution:

$$P_{n,ts}' = \frac{\delta_{ts}'^n}{n!} e^{-\delta_{ts}'}, 0 \leq n < \infty$$

(14)

C. The Calculation Steps

The establishment of the Monte Carlo simulation based on the Furman model helps to understand the entire secondary electron emission process and the formation of the energy spectrum in more detail. The Monte Carlo simulation steps of the Furman model can be divided into four steps:

step1: Record the energy of the incident electron E_0, the angle between the incident electron and the normal of the material surface θ_0 ;

step2: According to Furman's calculation formula, substitute the formula of 3 kinds of electrons to calculate the output yield of the 3 kinds of electrons;

step3: Using the Monte Carlo method, generate a random number of (0,1), obey the uniform distribution, and determine the number of secondary electrons generated;

step4: Returns to start the calculation of the next electron.

III. EXPERIMENTAL SIMULATION AND ANALYSIS

A. Energy Spectrum Curves at Several Incident Energies

The energy spectrum distribution of the secondary electron has essential research significance and has become an effective method for the analysis of the surface structure of the material and the study of surface physics. This article focuses on the relationship between the energy spectrum distribution of the secondary electron and the incident energy and angle of incidence.
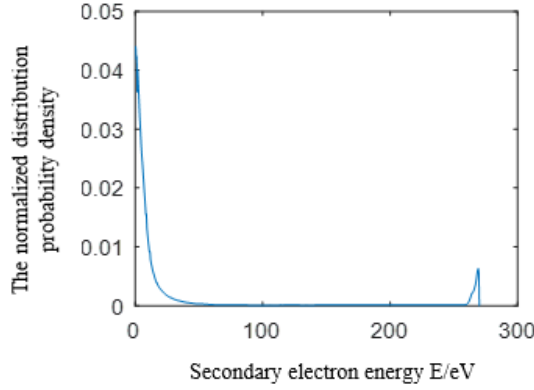
Fig. 2. The secondary electron energy spectrum with a sample number of 10000 and incident energy of 270 eV.
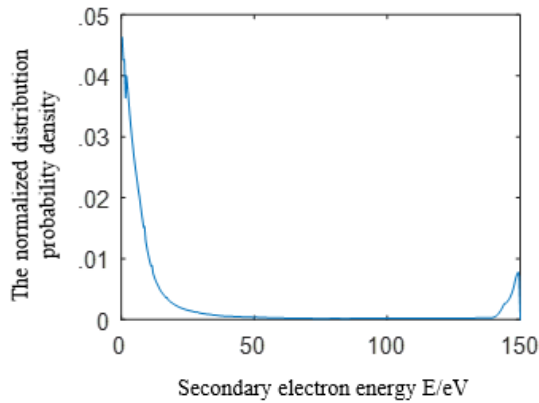
Fig. 3. Secondary electron energy spectrum with 20000 samples and 150 eV incident energy.

The numerical calculation results show the distribution characteristics introduced in Chapter 1. The reason why the number of samples is lower when the energy is low is that the program is based on a certain energy interval. The energy spectrum distribution curve here is normalized by area.
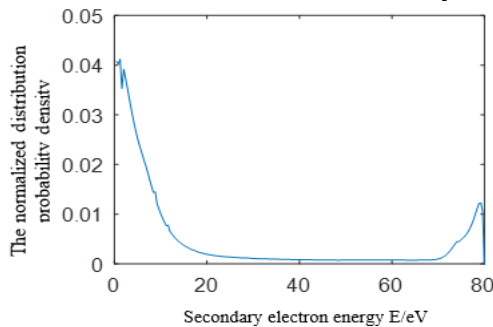
Fig. 4. Secondary electron energy spectrum with a sample number of 4000 and the incident energy of 80 eV.
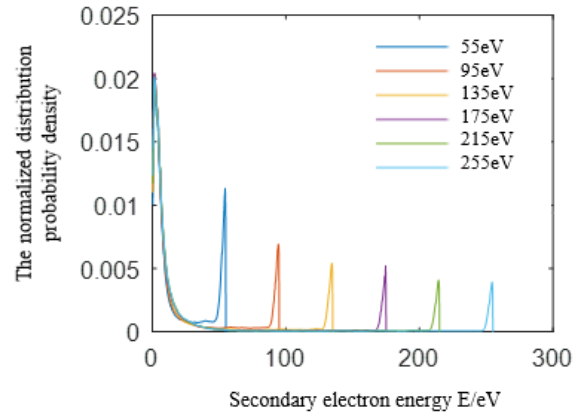
Fig. 5. Energy spectrum curves at several incident energies.

It can be observed that as the incident energy increases, the distribution probability density of secondary electrons in the high-energy part decreases, and the low-energy part increases. Combining this experiment and the characteristics of the secondary electron emission yield curve, analyze the reason for this problem: at low energy, since the energy of the incident electron is low, the internal secondary electrons generated are also less, so the secondary electron emission yield is relatively small, so the energy spectrum distribution presents the characteristics of a relatively small proportion of low-energy parts. When the energy increases, the internal secondary electrons excited inside the solid increase, and the total secondary electron emission yield increases. When the energy continues to increase, the incident depth of the incident electron becomes larger. Because the distance of the internal secondary electrons moving to the solid surface becomes longer, the distance becomes longer, and the energy loss becomes larger with the increase, so the secondary electron emission yield curve appears with the characteristic of increasing first and then decreasing, there is a maximum secondary electron emission yield.

B. Energy Spectrum Curves at Different Incident Angles

The incident angle of incident electrons has a relatively large impact on SEY. Therefore, in the process of numerical simulation research, the impact of angle changes on SEY must be considered.

Fig. 6. Energy spectrum curves at several incident angles.

Fig. 7. The relation curve between the secondary electron emission coefficient and the electron incident angle.

Fig. 8. Second electron emission factor.

can generate more internal secondary electrons in the future, so the emission yield of the secondary electrons becomes smaller.

C. Calculation and Comparison of Secondary Electron Emission Coefficient with Experimental Data

As shown in Fig.8, in the energy range of 160eV to 270eV, the secondary electron emission yield calculated by the model is in good agreement with the experimental results. At lower energies, there is a large error. This error may be due to insufficient Monte Carlo calculations and may be due to the test conditions of the experiment.

IV. CONCLUSIONS

This article focuses on the Monte Carlo simulation of the Furman model. The Monte Carlo simulation using the Furman model is used to calculate the secondary electron emission yield and compare it with the experimental results. In the energy range of 160eV to 270eV, the calculated secondary electron emission coefficient is in good agreement with the experimental results. The influence of incident energy and incident angle on the energy spectrum distribution of secondary electrons and the secondary electron emission coefficient was explored. The conclusion is that incident energy has a more significant influence on the energy spectrum distribution of secondary electron emission, and the incident angle has little influence on the energy spectrum of secondary electrons. The influence of the distribution is not apparent, and a theoretical analysis of the obtained results is made. It is found that the distribution density of the secondary electrons in the high energy part decreases with the increase of the incident energy and increases in the low energy part.

The secondary electron emission yield is the most commonly used physical quantity to characterize the secondary electron properties of materials. Therefore, the research results of this article exploring the influence of incident electron energy and incident electron angle on the secondary electron emission coefficient have important application significance.

As shown in Fig.6, the incident angle has less influence on the energy spectrum curve than the incident energy has on the energy spectrum curve.

As shown in Fig.7, the secondary electron emission yield increases slowly in the range of 0 to 30 degrees. As the angle increases, the increasing speed becomes faster and faster, and that's because when the angle goes up, the depth of the electron goes down, so more electrons can generate internal secondary electrons in the area with a smaller incident depth. These internal secondary electrons can escape to the surface of the material. Therefore, as the incident angle increases, the secondary electron emission coefficient increases. However, as the incident angle increases, the electrons will escape to the surface of the material before they

REFERENCES

[1] Miyake, Hiroaki, et al. "Secondary electron emission measurement of insulating materials for spacecraft." International Symposium on Discharges & Electrical Insulation in Vacuum IEEE, 2006.

[2] Clerc, S, et al. " On the computation of secondary electron emission models.." IEEE Transactions on Plasma ence.vol.34,pp:2219-2225,2006.

[3] Furman, M.A, and Pivi, M.T.F. " Probabilistic model for the simulation of secondary electron emission." Phys.rev.st Accel.beams vol.5,pp:317-322,2002

Theoretical Calculation of Photoemission Yield Spectrum of Space Metal Materials

Changxi Li
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
3120304022@stu.xjtu.edu.cn

Yu Chen*
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenyu@xjtu.edu.cn

Yan Yang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
yangyan0415@stu.xjtu.edu.cn

Zecai Chen
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenzecai@stu.xjtu.edu.cn

Hanzhi Li
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
lhz3119304518@xjtu.edu.cn

Shuang Wang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
shuang@xjtu.edu.cn

Abstract— **In the theoretical calculation of the photoemission yield spectrum of space metal materials, the mathematical model is found based upon three steps process of photoemission combined with the metal free-electron gas model. We take Cu as an example, the variation of photoemission yield spectrum caused by the electric field with different field strength and different incidence angle is calculated. This paper shows that the regions of quantum efficiency of Cu and Au when the materials are in the vacuum, under certain conditions, are respectively calculated between 8×10^{-8} and 2.6×10^{-3}, 8.53×10^{-6} and 3.8×10^{-3}, corresponding to the regions of incident ultraviolet wavelength are respectively between 138nm and 287nm,108nm and 249nm. Its quantum efficiency decreases with the increase of incident wavelength. The greater the intensity of the applied electric field, which is perpendicular to the copper surface, the greater the value of the calculated copper yield spectrum. The greater the angle formed by the incident ultraviolet light with the copper surface, the smaller the value of the calculated copper yield spectrum.**

Keywords—space metal materials, photoemission, yield spectrum

I. Introduction

For solving the problem of surface electrification caused by solar ultraviolet radiation during the orbit of the spacecraft, it is essential to study the factors affecting the surface electrification of the exposed materials of the spacecraft. In terms of space material properties, the photoelectron emission yield spectrum is an important parameter that affects the charging level. Metal is one of the important materials for spacecraft, and the calculation of its yield spectrum is of great significance for the simulation of the surface charging of the spacecraft and the overall design of the spacecraft. Therefore, this thesis carried out the numerical calculation and influence analysis of the photoelectron emission yield spectrum of metal materials under vacuum ultraviolet light irradiation.

This paper mainly discusses the three-step process of photoelectron emission proposed in the literature [1] and interprets it as a mathematical probability model, as well as the quantum efficiency calculation model proposed in the literature [2,3], which is based on the three-step process and derived by introducing the metal free-electron gas model. The efficiency calculation model is applied to the calculation of the photoelectron emission yield spectrum of copper and gold materials in the vacuum in this paper. We take the copper material as an example, and the influence of the applied electric field and incident angle on the photoelectron emission yield spectrum is analyzed.

II. Calculation Model Of Photoelectron Emission Yield Spectrum

A. Photoelectron Emission Yield Spectrum

The photoelectron emission yield spectrum is an important indicator for judging the difficulty of charging spacecraft materials, and its calculation method is the key research object of this paper. Therefore, we must first explain the definition of photoelectron and photoelectron emission yield spectrum.

The definition of photoelectron: when light radiates to the surface of the material, the electrons in the material absorb incident photons and become excited electrons, and the free electrons that escape the surface of the material become photoelectrons.

For the photoelectron emission yield spectrum of a material, the following definition needs to be made in this thesis: For the same material, under the same conditions, its surface is irradiated by the same wavelength of the incident light, and each incident photon can excite the surface of the material. The number of photoelectrons is the quantum efficiency of the material, and the function of quantum efficiency with different wavelengths (or with different incident photon energies) is the photoelectron emission yield spectrum. Its mathematical expression can be expressed as:

$$QE(\lambda) = \frac{n_e(\lambda)}{n_p(\lambda)} \quad (1)$$

where QE (λ) is the photoelectron emission yield spectrum of the material, n_e is the number of excited photoelectrons, n_p is the number of incident photons, and λ is the wavelength of the incident light. For the calculation of the photoelectron emission yield spectrum, the above definition is mostly used in the calculation of experimental results.

B. Three-step Process of Photoelectron Emission

For the calculation of the photoelectron emission yield spectrum, the above definition is mostly used in the calculation of experimental results, but for the theoretical calculation, it is unable to meet the requirements. Therefore, we need to start from the physical process of photoelectron emission and establish a mathematical model for theoretical calculations.

As for the physical process of light irradiating the surface of the material to make the electrons in the material escape

the surface of the material and become free electrons, the main model basis at this stage is the three-step photoelectron emission process proposed by Spicer [1] in 1958:

1) The process in which photons are incident on the surface of the material and are absorbed by the electrons in the material so that the electrons are excited;

2) The process of the excited electrons being transported to the surface of the material, and energy loss processes such as phonon-electron scattering and electron-electron scattering should be considered;

3) The process in which the excited electrons cross over the surface barrier and then escape from the surface of the material to become free electrons.

C. Calculation Model of Photoelectron Emission Yield Spectrum

Regarding the calculation of the photoelectron emission yield spectrum, the three-step model of photoemission by the physical process of photoemission can be considered to be converted into a probability problem. The quantum efficiency the probability of an incident photon causing the electrons in the material to absorb the photons and becoming excited electrons, which can maintain sufficient energy during the process of transporting to the material surface, and finally can cross over the material surface barrier and become free electrons. The function of quantum efficiency with respect to wavelength is the photoelectron emission yield spectrum. Therefore, based on the above analysis, it can be concluded that the photoelectron emission quantum efficiency can be converted into the probability of an electron experiencing all these three steps in the photoelectric emission process. According to the analysis, the following mathematical probability model is derived:

$$QE = P_1 \times P_2 \times P_3 \qquad (2)$$

where QE represents the quantum efficiency of photoelectron emission, P_1 represents the probability that electrons in the material absorb incident photons and become excited electrons; P_2 represents the probability that the excited electrons can maintain sufficient energy during the process of transport to the material surface; P_3 represents the probability that excited electrons can escape from the surface of the material and eventually become free electrons.

III. NUMERICAL CALCULATION OF PHOTOELECTRON EMISSION YIELD SPECTRUM OF SPACE METAL MATERIALS

A. Calculation Model of Metal Photoelectron Emission Yield Spectrum

The mathematical model for calculating the quantum efficiency of the photoelectric emission of the material is obtained above. For the derivation of the photoelectron emission yield spectrum equation, the mathematical model for calculating the quantum efficiency needs to be derived as a function of the wavelength change of the incident light according to the specific material.

Here, from the discussion and derivation in the literature [2,3], it can be known that the establishment of the mathematical probability model of the photoelectron emission yield spectrum of space metal materials is based on the theoretical model of metal free-electron gas and also considers the Fermi-Dirac electron statistical distribution, the

electron-electron scattering process, and the electrons at the interface of the metal and vacuum The effect of the exit angle on the photoelectron emission, and through some reasonable assumptions and simplifications, the calculation expression of the photoelectron emission yield spectrum of space metal materials is finally deduced. The Eq.(2) can be derived as the following Eq.(3)

$$QE = \frac{1-R}{1+\dfrac{\lambda_{opt}}{2\lambda_{e-e}(E_m)}\dfrac{\hbar\omega\sqrt{\phi}}{E_m^{3/2}}(1+\sqrt{\dfrac{\phi}{\hbar\omega}})} \frac{(E_F+\hbar\omega)}{2\hbar\omega}[1-\sqrt{\dfrac{E_F+\phi}{E_F+\hbar\omega}}]^2 \qquad (3)$$

Where QE represents the quantum efficiency, R represents the reflectivity of the material, λ_{opt} represents the incident depth of light, E_m represents specific energy above the Fermi level, ϕ represents work function, and E_F represents the Fermi level of the metal material. The above equation is used to calculate the yield spectrum of space metal materials in this paper.

B. Numerical Calculation of Yield Spectrum of Space Metal Materials

Here, copper materials and gold materials are selected as the metal materials for this photoelectron emission yield spectrum calculation. Through recent reports and materials, the parameter values used in this calculation are shown in Table 1 below:

TABEL I. THE CALCULATION PARAMETERS OF THE PHOTOELECTRON EMISSION YIELD SPECTRUM OF COPPER AND GOLD MATERIALS:

Parameters	Cu	Au
E_F	7.0eV	5.53eV
ϕ	4.31eV	4.81eV
R	[4]	[5][6]
λ_{opt}	[4]	[5][6]
E_m	8.6eV	8.6eV
λ_{e-e}	2.2nm	2.2nm

For the calculation of the photoelectron emission yield spectrum of gold and copper materials, since the calculations are made for space metal materials, the external medium is the vacuum, considering that the external electric field of the material is 0MV/m, the temperature is much lower than room temperature, and the incident wavelength is considered to be the ultraviolet band. Therefore, the photon energy range of the incident light is selected to be 4eV-12eV (corresponding to the wavelength range of the incident light is 100nm-300nm), and the angle between the incident light and the material surface is 90 degrees, that is, vertical incidence. Under such external conditions, the calculation parameters of the two metal materials referred to in Table 1 are brought into Eq.3, and the software MATLAB is used to program the calculation and draw the calculation curve, as shown in Figure 1.

The calculation results indicate that the regions of quantum efficiency of Cu and Au when the materials are in the vacuum, under certain conditions, are respectively calculated between 8×10^{-8} and 2.6×10^{-3}, 8.53×10^{-6} and 3.8×10^{-3}, corresponding to the regions of incident ultraviolet wavelength are respectively between 138nm and

287nm,108nm and 249nm, and its quantum efficiency decreases with the increase of incident wavelength.
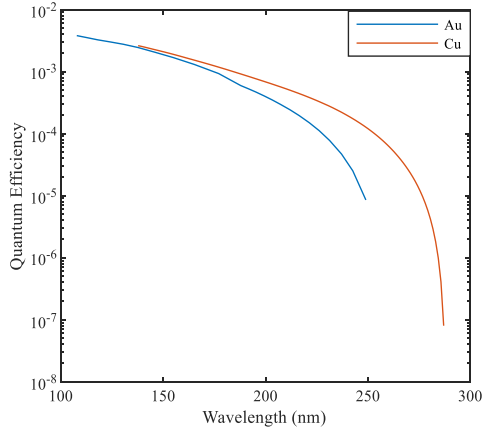
（a）Calculation curve of yield spectrum with wavelength as horizontal axis
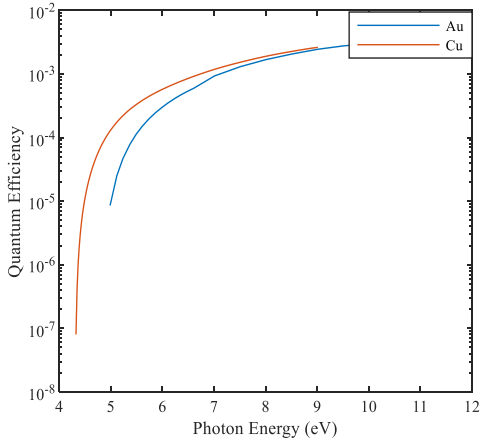
（b）Yield spectrum curve with photon energy as the horizontal axis

Fig. 1. Calculation curve of photoelectron emission yield spectrum of copper and gold

IV. ANALYSIS OF THE INFLUENCE OF SPACE METAL YIELD SPECTRUM CALCULATION

A. Analysis of the Influence of External Electric Field on Calculation of Space Metal Yield Spectrum

For the influence of the external electric field on the photoelectron emission yield spectrum of metal materials, it is necessary to first consider the potential barrier on the metal surface in more detail. According to Schottky auxiliary electric field emission theory, considering the influence of Coulomb force and applied electric field, the potential barrier on the metal surface can be expressed by the following Eq. 4:

$$\phi_{eff} = \phi - \phi_{\text{schottky}} = \phi - \left(\frac{e^3 E_x}{4\pi\varepsilon_0} \right)^{1/2}$$

(4)

Where, ϕ_{eff} represents the effective work function that takes into account the influence of the mirror image force caused by the electron leaving the metal surface and the effect of the applied electric field, ϕ is the original work function, and ϕ_{schottky} is the Schottky work function mentioned above. E_x is the electric field strength of the externally applied electric field, e is the element charge of

1.6×10^{-19}C, ε_0 is the vacuum dielectric constant of 8.8542×10^{-12}F/m.
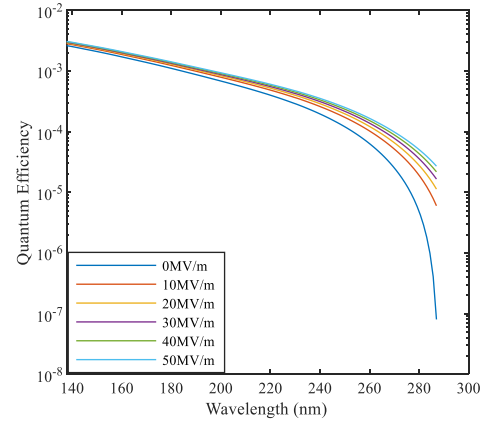
Fig. 2. Photoelectron emission yield spectrum curves of copper materials under external electric fields with different field strengths

As shown in Figure 2, the photoelectron emission yield spectrum curve of the copper material shows an overall upward trend with the increase of the applied electric field strength. This verifies that under certain external conditions, the field strength of the same metal material increases from 0MV/m to 50MV/m with the applied external electric field at the same incident wavelength, and known from Schottky auxiliary electric field emission theory, the barrier height of the metal surface continues to decrease. As a result, the work function continues to decrease, and the excited electrons that absorb photons inside the metal material are more likely to escape from the surface and reach the vacuum to become free electrons, thereby increasing the quantum efficiency of the material.

B. Analysis of the Influence of Incident Angle on Calculation of Space Metal Yield Spectrum

The incident angle of the light will affect the probability of the light being reflected when it radiates to the metal surface, thus affecting the final calculation result of its yield spectrum. First, use the Fresnel equation to find the polarization direction perpendicular to and parallel to the plane composed of the light incident direction and the normal line of the metal surface, as shown in Eq. 5:

$$\begin{cases} R_s = \left(\dfrac{n_1 \cos\theta_1 - \sqrt{n_2^2 - n_1^2 \sin^2\theta_1}}{n_1 \cos\theta_1 + \sqrt{n_2^2 - n_1^2 \sin^2\theta_1}} \right)^2 \\ R_p = \left(\dfrac{n_2^2 \cos\theta_1 - n_1\sqrt{n_2^2 - n_1^2 \sin^2\theta_1}}{n_2^2 \cos\theta_1 + n_1\sqrt{n_2^2 - n_1^2 \sin^2\theta_1}} \right)^2 \end{cases}$$

(5)

Where R_s is the reflectivity of the polarization direction perpendicular to the surface, and R_p is the reflectivity of the polarization direction parallel to the surface. n_1 is the refractive index of the original medium, n_2 is the refractive index of the incident medium, and θ_1 is the incident angle.

At the same time, since the calculated light source is the ultraviolet band in the sunlight, it is natural light, so the light intensity in each polarization direction is the same, so the total reflectance calculation is shown in Eq.6:

$$R = \frac{R_s + R_p}{2}$$

(6)

At the same time, since the original medium where the incident light is located in the vacuum, the refractive index n_1 can be set to 1. For the incident metal material, copper material is selected. Due to the need to consider its light absorption properties, its refractive index n_2 is a complex refractive index, $n_2=n+ik$, where n is the real refractive index of the material, and k is the extinction coefficient of the material, and both values will vary with the material The wavelength of different incident light changes, which in turn affects the reflectivity of the material. Secondly, the Eq.5 for calculating the complex refractive index is still applicable and only needs to be regarded as the square of the complex modulus of the final calculation result.
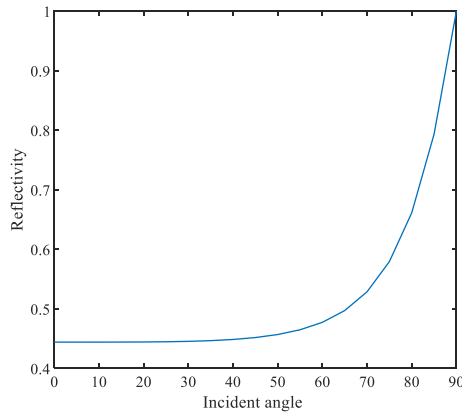
Fig. 3. The reflectivity of the copper material vs. the incident angle when the incident wavelength is 220nm
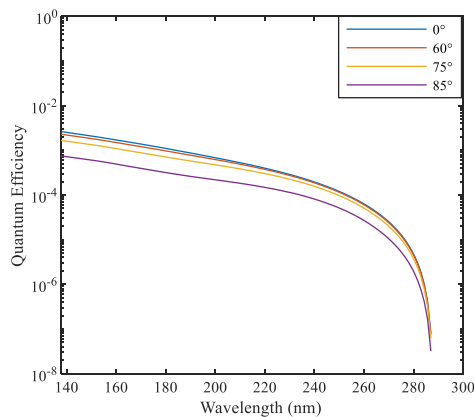
Fig. 4. Variation of photoelectron emission yield spectrum of copper at different incident angles

For the same copper material in the vacuum, the same wavelength (that is, the same complex refractive index, select the complex refractive index n=1.18119, k=1.93297 of copper at a wavelength of 220nm), the reflectivity change caused by combining Eq.5 and Eq.6, use MATLAB programming calculation and drawing; for the same material in the vacuum, different wavelengths (different complex refractive index), the changes in the photoelectron emission yield spectrum of metal materials caused by the combination of Eq.5, 6 and 3, and the remaining parameters are determined as above consistent, use MATLAB programming

to calculate and plot. The calculation and drawing results are shown in the figure above.

It can be seen from Fig. 3 that the increase in incident angle leads to a sharp rise in reflectivity in the region greater than 60 degrees, which causes a rapid decline in the photoelectron emission yield spectrum in Fig.4 in the same region. It can be seen from the Fig.4 that as the incident angle increases, the quantum efficiency of copper has a decreasing trend at the same wavelengths, and the rate of decrease is also getting faster. The physical meaning of the calculation result is: as the angle of incidence increases, the reflectivity shows an upward trend, the absorption rate of the incident photon by the copper material decreases, and the probability that the valence electrons inside the copper material which can absorb photons and become excited electrons decreases. This causes a decline in the photoelectron emission yield spectrum curve as the incident angle increases.

V. CONCLUSION

This paper indicates that the regions of quantum efficiency of Cu and Au when the materials are in the vacuum, under certain conditions, are respectively calculated between 8×10^{-8} and 2.6×10^{-3}, 8.53×10^{-6} and 3.8×10^{-3}, corresponding to the regions of incident ultraviolet wavelength are respectively between 138nm and 287nm,108nm and 249nm., and its quantum efficiency decreases with the increase of incident wavelength. The applied electric field perpendicular to the copper surface, the greater the intensity of the electric field, the greater the value of the calculated copper yield spectrum; the greater the angle formed by the incident ultraviolet light with the copper surface, the smaller the value of the calculated copper yield spectrum.

REFERENCES

[1] W. E. Spicer, "Photoemissive, photoconductive, and optical absorption studies of alkali-antimony compounds," Physical Review, vol. 12, no. 2,pp.117-122,1958

[2] D. H. Dowell, F. K. King, R. E. Kirby, and J. F. Schmerge, "In situ cleaning of metal cathodes using a hydrogen ion beam," Physical Review Special Topics, vol. 9, no. 6,pp .0635021-0635028 ,2006

[3] David H. Dowell and John F. Schmerge, "Quantum efficiency and thermal emittance of metal photocathodes," Physical Review Special Topics, vol. 12, no. 7,pp.0742011-07420110 ,2009

[4] K. Stahrenberg, Th. Herrmann, K. Wilmers, N. Esser, W. Richter, and M. J. G. Lee, "Optical properties of copper and silver in the energy range 2.5-9.0 eV,"Physical Review. vol. 64,no. 11,pp.1013-1092,2001

[5] P. B. Johnson and R. W. Christy, "Optical constants of the noble metals," Physical Review, vol. 6, no. 12,pp. 4370-4379, 1972

[6] W.S.M.Werner, K.Glantschnig and C.Ambrosch-Draxl., "Optical constants and inelastic electron-scattering data for 17 elemental metals," The Journal of Chemical Physics Review, vol. 38, no. 4, pp. 1013-1092,2009

[7] W. F. Krolikowski and W. E. Spicer, "Photoemission studies of the noble metals I. Copper," Physical Review.vol. 1, no. 2,pp.882-887,1970

[8] W. F. Krolikowski and W. E. Spicer, "Photoemission studies of the noble metals. II. Gold". Physical Review, vol. 1,no. 2,pp.479-481,1970

[9] C. N. Berglund and W. E. Spicer, "Photoemission studies of co pper and silver: theory," Physics Review, vol. 136, no. 4A,pp.1 044-1048,1964

[10] C. N. Berglund and W. E. Spicer, "Photoemission studies of co pper and silver: experiment," Physics Review, vol. 136, no. 4A, pp.1031-1039,1964

Development of a Measurement System for the Secondary Electron Emission Yield Spectrum of Space Materials

Yu Chen*
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenyu@xjtu.edu.cn

Guorui Huang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
huang0923@stu.xjtu.edu.cn

Yan Yang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
yangyan0415@stu.xjtu.edu.cn

Qingyun Shi
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
sqy3119304516@stu.xjtu.edu.cn

Zecai Chen
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenzecai@stu.xjtu.edu.cn

Shuang Wang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
swang_helene@foxmail.com

Yonghong Cheng
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
cyh@xjtu.edu.cn

Abstract—When a spacecraft is operating in space orbit, the phenomenon of secondary electron emission will occur under the influence of charged particles in space, causing the surface of the spacecraft to charge, and long-term accumulation will cause damage to the spacecraft and its instruments. In response to the needs of spacecraft live protection and life evaluation, this paper has researched the secondary electron emission yield spectrum test technology of space materials and completed the development of the test system. Firstly, we use an electron gun with an energy range of 0-2000eV, a beam current of 1nA-10μA, and a minimum beam spot diameter of 0.5mm to achieve stable electron beam current in the secondary electron emission coefficient measurement system in this paper. Secondly, the vacuum system uses mechanical pumps, molecular pumps, and sputtering ion pump, so that it can reach a vacuum of 1×10^{-6}pa. Thirdly, the measurement scheme of vacuum secondary electron collecting ball and grid structure is proposed for the first time, which can realize the distinguished collection and measurement of true secondary electrons and backscattered secondary electrons. Finally, the measurement of the secondary electron yield spectrum of metal materials was studied, and the efficient measurement of the secondary electron emission yield spectrum of oxygen-free copper and gold materials was completed. The test results have reached that the secondary electron emission yield spectrum data of Cu and Au samples under 0-2000eV incident electrons is consistent with the literature, which verifies the reliability of the test system.

Keywords—Secondary electron yield spectrum，Measurement

I. Introduction

Spacecraft now play a vital role in our military's defense capabilities and modern living standards and have been widely used in satellite broadcasting, satellite communications, and scientific research. In recent years, there have been endless reports of spacecraft and satellite failures caused by surface or internal charging. According to the 6th Spacecraft Charging Technology Conference, electrostatic discharge accounts for 54% of all spacecraft accidents. SEU accounts for 29%, radiation damage accounts for 12%, and others account for 5%[2]. The main reason which causes the spacecraft faults, that is electrostatic and discharge.

The spacecraft charging effect is divided into surface charging and internal charging. It is one of the leading causes of spacecraft abnormalities and failures caused by the space environment. Surface charging is the earliest space environmental effect discovered by people, and secondary electron emission affects the surface. Given the above reasons, it is very urgent to observe and study the radiation charging of spacecraft. Studying the secondary electron emission coefficients of insulators and conductors in the space environment will reduce the incidence of spacecraft and equipment failures and improve the safe operation of the spacecraft.

The measurement technology of the secondary electron yield spectrum of metal materials is relatively mature. The current method and collection method can be used to obtain accurate and reproducible secondary electron yield spectrum data. But for insulating materials, the electron beam of the electron gun will cause charge accumulation on the surface of the insulating material. Excessive charge accumulation will affect the measurement of the secondary electron emission coefficient. Thomson[2] used a method of bombarding an area of about 1~10mm^2 on the sample with a low current beam in a short time to minimize the charge dislocation. Johnson[3] used a secondary electron coefficient measurement system with a heating device to study for solid inorganic materials. Besides, it also used a charge neutralization method, which is very The surface with positive potential is neutralized with a low-energy electron flow gun. In contrast, the sample with negative potential is neutralized with a variety of visible light and ultraviolet light. To overcome the constant charge deposition on the insulating material, we minimized the current size and the time for the electron beam to irradiate the sample, so the low current pulse scanning method is used to measure the secondary electron emission coefficient of the insulating material.

*Yu Chen is the corresponding author. (e-mail: chenyu@xjtu.edu.cn).

This paper will mainly introduce the secondary electron emission yield spectrum measurement system designed and developed by our research team.

II. THEORETICAL RESEARCH

The sample atoms and external electrons that are bombarded by the incident electron beam and leave the sample surface are called secondary electrons. This kind of electron is a free electron in a vacuum. Since the bonding energy between the nucleus and the outer valence electrons is minimal, the outer electrons are a higher probability to separate from the atoms and ionize the atoms. When an incident electron with high energy is injected into the sample, it can generate many free electrons. 90% of these free electrons are valence electrons from the outer layer of the sample atoms. The energy of secondary electrons is relatively low, generally not more than 50eV. Most secondary electrons have only a few electron volts of energy. Secondary electrons are generally emitted in the depth range of 5-10nm in the surface layer. They are very sensitive to the surface morphology of the sample. Therefore, it can display the surface morphology of the sample very effectively.

When an electron bombards the surface of the material, electron emission may occur, depending on the energy of the electron. If the energy is lower than 10eV, electrons may be reflected from the surface. High-energy electrons interact with neighboring electrons in the surface material, and energy is thus dispersed. One or more of the neighboring electrons may leave the surface due to obtaining sufficient energy as a secondary electron. That is, incident electrons (primary electrons) bombard one or more electrons from the surface of the material, which is secondary electrons.
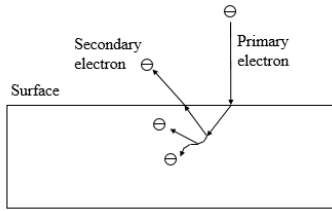
Fig. 1. Example of secondary electron emission

When electrons or ions bombard a substance, the ratio of the number of electrons (secondary electrons) leaving the surface of the substance to the number of electrons (primary electrons) bombarding the substance is called the secondary electron emission coefficient δ of the substance. The formula can express a true secondary electron δ.

$$\delta = \frac{I_s}{I_p} \tag{1}$$

In the formula, I_s is the true secondary electron current, and I_p is the primary electron current. And the total electron σ can be expressed as:

$$\sigma = \frac{I_s + I_{bse}}{I_p} = \frac{I_A}{I_A + I_{coll}} \tag{2}$$

I_{bse} is the backscattered electron current, I_A is the current on the sample, I_{coll} is the current on the collecting ball. The backscattered electron can be expressed as η.

$$\sigma = \delta + \eta \tag{3}$$

After the primary electron adds energy to the electrons in the solid, not all the energy-obtained electrons can leave the surface. Those emitted electrons are called secondary electrons. When a primary electron bombards the surface of the material, the probability of the backscattered electron is η, which usually occurs at some ion positions. The emitted electrons are the same as the incident electrons. The probability η depends on the material, the energy of the primary electron, and the angle of incidence. Unlike secondary electrons, η does not exceed 1, and the amount of energy of backscattered electrons is almost close to the energy of the primary electron.

The sum of the secondary electrons and the backscattered electrons is the total electron emission probability. For an incident electron with energy E, there are $\delta + \sigma$ outgoing electrons. There are many incident electrons with different energies. For an incident electron current I_p, there are two outgoing electron currents. Namely, the secondary electron current I_s and the backscattered electron current I_{bse}. The incident current is defined as the number of ions (or the amount of charge) that arrive in a specific area per unit time.

III. SYSTEM DEVELOPMENT

This test system mainly controls the electron gun and oscilloscope through a pulse generator colleague. It obtains the incident electron beam and the secondary electron current waveform of the collector within 30 microseconds pulse width, to read the waveform amplitude and calculate the secondary electron emission coefficient of the insulating material.
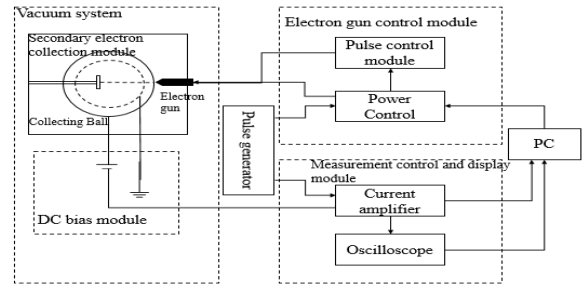
Fig. 2. Measurement scheme of secondary electron emission coefficient

The system is mainly divided into secondary electron collection modules (Collecting ball, inner grid, outer grid, sample), electron gun and its control module (electron gun, control power supply, pulse generator), measurement control and display module (oscilloscope, Four parts: variable gain high-speed low-current amplifier) and PC. This system can achieve automatic control and measurement of the secondary electron emission coefficient.
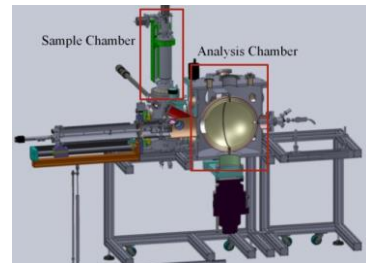
Fig. 3. Physical map of secondary electron yield spectrum measurement system

The sample to be tested transferred from the sample chamber to the analysis chamber is placed and fixed on the four-dimensional sample stage. The sample stage adopts a high-precision ultra-high vacuum design and is driven by both

manual and stepping motors. The analysis chamber heats the entire sample table through a sample heating mechanism, with a maximum temperature of 120°C, to complete the photoelectron emission coefficient and secondary electron emission coefficient tests at different temperatures. The sample stage can be driven by ultra-high vacuum linear, up and down, left and right, and inclination through a four-dimensional sample drive mechanism. In this way, it can realize the scanning test of the photoelectron emission coefficient and the secondary electron emission coefficient of the sample. The XY drive range of the four-dimensional sample stage is ±15mm, the Z drive range is not less than 400mm, and the tilt drive range is 0-85°.

A. Vacuum System

The vacuum test system is used to perform secondary electron emission yield spectra of solid materials such as metals, semiconductors, and insulators. Not only can the secondary electron emission coefficients under different energies and different angles be obtained, but also the secondary electrons under each condition can be obtained. Energy distribution. The whole set of equipment is composed of sample chamber, analysis chamber, and electron gun, supplemented by corresponding components such as vacuum acquisition, vacuum detection, sample transfer, signal acquisition, and power supply.

The function of the sample chamber is to realize the heating and degassing pretreatment of the sample and the sample. The analysis chamber is to measure the secondary electron of solid materials such as metals, semiconductors, and insulators by collecting the secondary electrons emitted by the sample. At the same time, the surface potential of the sample can be measured. The surface potential test probe interface is reserved in the vacuum chamber.

An ultra-high vacuum gate valve is installed between the analysis chamber and the sampling chamber. When the gate valve is closed, the two chambers are independent of each other, and when the gate valve is opened, the two chambers are connected and can carry out sample transfer. The whole system is entirely independently designed and manufactured, with convenient operation and reliable performance.

B. Collection Device

There are three signals to measure the secondary electron emission coefficient, one is the intensity of the initial incident electron beam, the other is current when the electron irradiates the sample, and the third is the secondary electron current on the secondary electron collector. The design of the sample stage can obtain the initial incident electronic signal and target current signal. To obtain accurate secondary electron emission coefficient measurement results, the structural design of the secondary electron collection assembly is the key.
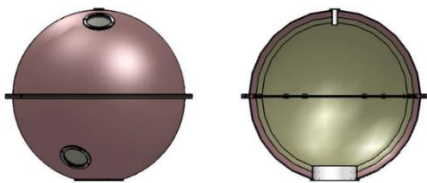
Fig. 4. Collecting ball

The secondary electron collection combination is composed of two layers of grids and a layer of the metal collector. The first layer of the grid is grounded to shield the external electric field; the second layer of the grid can be applied with a voltage. The charge signal obtained by the collector is detected by connecting the coaxial electrode with an external instrument. The potential between the electrodes of each layer is different. And the insulation and wiring between the electrodes of each layer are concentrated on the equatorial plane.

C. Hardware Selection

1) Electron Gun

The ELG-2/EGPS-1022 electron gun from KPI of the United States is used as the electron beam source for testing the secondary electron emission coefficient of space materials. It is a complete subsystem.

Fig. 5. KPI ELG-2 / EGPS-1022 electron gun

Among them, ELG-2 is an electron gun, with the EGPS-1022 control power supply, it can generate a pulsed electron beam with an energy of 1~2000eV and a pulse width of 20ns~500μs. The control unit allows fast condenser beam pulses or optional double gate pulses. For capacitive pulses, using a pulse connection box, fast electron beam pulses can reach 20 ns with a duty cycle of 20%. Capacitive pulse requires an external pulse signal generator. For double-gate pulses, only TTL pulse input is required.

The electron gun is equipped with special software, which can control the electron gun emission current, electron energy, first anode voltage, delete voltage, and focus voltage through a friendly user interface to realize the remote control of the electron gun.

2) Pulse Generator

The pulse control adopts the DG645 digital delay/pulse generator developed by the American SRS company. It is a multifunctional digital delay/pulse generator that can accurately define pulses with repetition frequency within 10MHz. Compared with the previous design, the instrument has been improved in many aspects, mainly including lower jitter, higher accuracy, faster trigger rate, and more output ports. DG645 interfaces include Ethernet, GPIB, and RS232 interfaces for the computer or network control.

3) Measuring Device

According to the requirements of nA-level micro-current measurement in the secondary electron test, the DHPCA-100 variable gain high-speed low-current amplifier produced by Germany FEMTO was selected. The gain is adjustable from 1×10^2 to 1×10^8V/A; the bandwidth is from DC to 200MHz; the upper cut-off frequency can be adjusted to 1MHz, 10MHz, or full bandwidth; AC/DC coupling can be switched; the bias voltage can be adjusted for external use photoelectric detector; input protection for ±1.5kV transient voltage; local and remote control of all main functions.

Especially for metal materials, we use Keithley's 6517b high-precision ammeter to measure current signals on the collecting ball, sample stage, and grid. The 6517b has a reading rate of 425 readings per second, which is significantly faster than the same type of electrometer and can provide a quick and easy way to measure weak currents.

IV. EXPERIMENTAL STEPS AND RESULTS

For metal materials, an ultra-high vacuum environment must be maintained during the metal surface analysis process. According to statistics, under a high vacuum of 7.5×10^{-7} Torr, a layer of gas is adsorbed on the solid surface every second, so it must be in an ultra-high vacuum environment to keep the sample surface clean during the analysis. Therefore, for the secondary electron emission coefficient test of the metal sample, it is necessary to use an argon ion gun for cleaning.

The secondary electron emission yield spectrum of copper is shown in Fig.6.
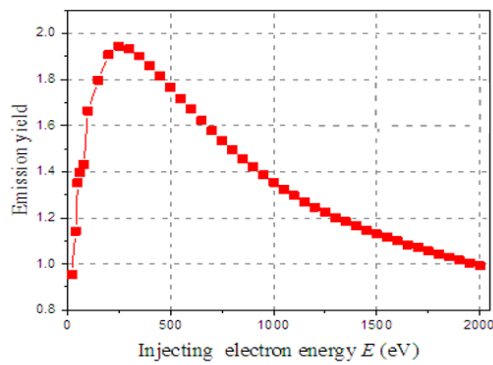
Fig. 6. Secondary electron emission coefficient of Oxygen-free Copper

From the figure, we can see that the secondary electron emission coefficient of the copper sample reaches a peak value of 1.94 at 250eV, and the zero-crossing points are at 20eV and 1950eV. Compared with the current copper secondary electron emission yield spectrum data measured by domestic and foreign laboratories, the result is accurate, which verifies the reliability of the measurement system. The secondary electron emission yield spectrum of Au is shown in Fig.7.
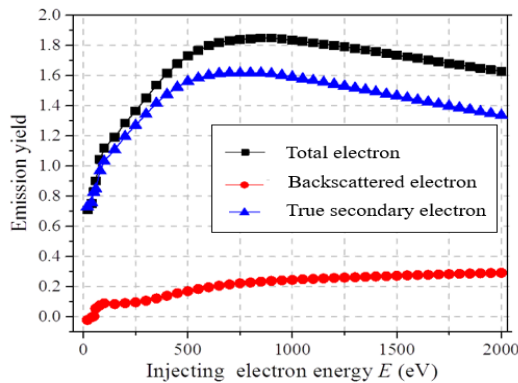
Fig. 7. Secondary electron emission coefficient of Au under different energy

The maximum total secondary electron emission coefficients obtained are all slightly higher than 1.8, and the peak incident energy of the corresponding curves are all around 600 eV.
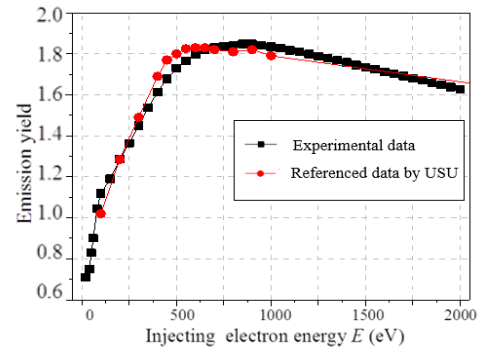
Fig. 8. Secondary electron emission coefficient of Au

The experimental SEY result of Au film is shown in Fig. 8, and it is compared with the referenced data by Utah State University (USU)[7]. From the figure, we can see that the secondary electron emission coefficient curve of the gold material characterized by the secondary electron test system developed in this paper is very close to the test curve given by USU.

V. CONCLUSION

Based on the analysis of the principle of secondary electron emission in the space environment, a measurement system of space material secondary electron emission yield spectrum was developed. The secondary electron yield spectrum of the metal material is tested by the direct current method to obtain the secondary electron yield spectrum of Cu and Au. The comparison with USU data verifies the accuracy and reliability of the secondary electron emission yield spectrum measurement of metallic materials in this paper. And pave the way for the subsequent measurement of insulating materials by the measurement system. The results of this paper are of great significance for the measurement of the secondary electron emission coefficient of spacecraft materials and the evaluation of material discharge.

REFERENCES

[1] Chen Y, Kouno T, Toyoda K, "Total electron emission yield measurement of insulator by a scanning small detector," Applied Physics Letters, 2011, 99(15):251.

[2] Koons H C, Mazur J E, Selesnick R S, Blake J B, Fennell J F, Roeder J L, Anderson P C, The Impact of The Space Environment on Space Systems, 6th Spacecraft Charging Technology Conference, 2000.

[3] Thomson C D, Zavyalov V, Dennison J R, "Instrumentation for studies of electron emission and charging from insulators," Spacecraft Charging Technology Conference. 2004.

[4] Johnson J B, "Secondary electron emission from targets of barium-strontium oxide," Physical Review, 1948, 73(9):1058-1073.

[5] Miao G H , Cui W Z , Yang J, "Research and development of the measurement equipment of secondary electron emission," Space Electronic Technology, 2018.

[6] Gross B , Hessel R , "Electron emission from electron-irradiated dielectrics," IEEE Transactions on Electrical Insulation, 1991, 26(1):18-25.

[7] Dennison J R., Sim A, Thomson C D, "Evolution of the electron yield curves of insulators as a function of impinging electron fluence and energy," IEEE Transactions on Plasma Science. 2006, 34(5): 2204–2218

[8] Chen Y, Wu J, Kunimitsu R, Gray A, Toyoda K, Cho M G, "Total electron emission yield of solar cell coverglass and optical solar Reflector," IEEE Transactions on Plasma Science. 2013, 41(12):3558-3564

In-pipe Detection System Based on Magnetic Flux Leakage and Eddy Current Detection

Xiaolin Liu
School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
liuxiaolin@buaa.edu.cn

Chun Hu*
School of Electronics and Information Engineering
Beihang University
Beijing, China
buaa_hc@buaa.edu.cn

Peng Peng
Service Robot Detection Department
Chongqing Dexin Robot Testing Center Limited Company
Chongqing, China
272508585@qq.com

Rui Li
PetroChina Pipeline Company
Langfang, China
kjlirui@petrochina.com.cn

Xiaoming Zhao
PetroChina Pipeline Company
Langfang, China
xmzhao@petrochina.com.cn

DeZhi Zheng
Innovation Institute of Frontier Science and Technology,School of Instrumentation and Optoelectronic Engineering
Beihang University
Beijing, China
zhengdezhi@buaa.edu.cn

Abstract—Pipeline plays an important role in the transportation of oil and gas. In-pipe detection technology is currently widely used at home and abroad. It is of great significance to develop in-pipe detection equipment with independent intellectual property rights and suitable for the actual situation of oil and gas pipelines in China. ANSYS software is used to simulate and analyze the leakage magnetic field of pipeline defect, and the influence rule of defect size on the leakage magnetic field is explored. An in-pipe detection system based on magnetic flux leakage and eddy current detection is designed and implemented, which takes the magnetic flux leakage (MFL) detection and eddy current (EC) detection as the main detection means, mainly including detection sensor and data acquisition and storage system. The system is carried by in-pipe detection equipment to carry out a pull test in a specially-made pipe. The results show that the system has a good detection effect for various types of defects, and the results of the pull test further verify the simulation results. The in-pipe detection system combines the advantages of MFL and EC detection, and can meet the needs of the pipeline detection tasks in China.

Keywords—in-pipe detection, magnetic flux leakage detection, eddy current detection

I. INTRODUCTION

Pipeline plays an important role in the transportation of oil and gas [1]. At the same time, pipeline leakage and other accidents also pose a great threat to national property and people's life safety. According to the law of our country, the main oil and gas pipelines need to be inspected every 3-5 years to evaluate and maintain the operation safety of the pipelines. In the pipeline detection methods, the internal detection technology can realize large-area, automatic pipeline detection, evaluate the safety status of the pipeline, and effectively avoid the occurrence of pipeline accidents.

Since the 1990s, pipeline integrity management has been carried out in foreign countries. From the pipeline commissioning stage, the internal detection technology has been applied to continuously monitor the pipeline status [2]. However, the engineering application of in-pipe detection technology in China started late, and there is a certain gap with foreign technology level. At present, most of the internal detection work of domestic pipelines is undertaken by foreign pipeline detection companies, and the foreign detection

equipment are equipped with navigation and positioning system. In the process of pipeline detection, important data related to national security such as oil and gas pipeline coordinates are collected and stored, which will pose a major threat to the national economic lifeline [3]. Therefore, it is of great significance to develop in-line detection equipment with China's independent intellectual property rights and suitable for the actual situation of China's oil and gas pipelines [4].

In-pipe detection technology can send pipeline detection equipment without stopping the transmission of medium. The equipment moves with the medium under the action of medium to realize the detection of pipeline. Due to the dynamic detection nature of internal detection technology, magnetic flux leakage (MFL) detection, geometric deformation detection, ultrasonic detection are widely used in pipeline internal detector, and eddy current (EC) method, magnetic memory method, weak magnetic method and other emerging technologies are gradually developed [5]. However, each detection technology has its own advantages and disadvantages and scope of application. Therefore, it is one of the development trends in the field of pipeline detection to integrate multiple detection technologies and develop equipment that can realize multiple detection functions.

The in-pipe detection system designed and implemented in this paper takes MFL detection technology as the core and EC detection technology as an auxiliary to realize the detection of defects on the inner and outer walls of the pipeline. It also includes a dedicated data acquisition and storage system. The ability of the system to detect defects is verified by the pull test in a specially designed pipeline with artificial defects. The results of the pull test show that the system has a good detection effect for various types of defects, and can meet most conventional pipeline detection tasks.

II. IN-PIPE DETECTION TECHNOLOGY

The in-pipe detection equipment is used to carry the internal detection system, which generally includes the power part, the detection sensor part, the data acquisition part and the battery part [6]. The power part provides the power for the equipment to move forward in the pipeline, generally relying on the pressure of the pipeline conveying medium to push forward. The detection sensor part is the core of the internal detection equipment, responsible for the detection of the pipeline, generally in the form of multiple probes. The data

---

* Chun Hu is the corresponding author. (e-mail: buaa_hc@buaa.edu.cn).

collection part undertakes the functions of data collection, processing and storage during the detection process. The battery part provides power for the normal operation of each part of the internal detection equipment.

A. Magnetic Flux Leakage Detection

MFL detection is the longest researched, most widely used and most mature technology. This technology is suitable for the identification and quantification of metal pipeline surface and internal defects. It has simple operation and high reliability, and has no special requirements for the pipeline inspection environment.

MFL detection technology is based on the feature of high permeability of ferromagnetic materials. After the ferromagnetic tube wall is magnetized by the magnetizing device, if the material of the tube wall is continuous and uniform, the magnetic field lines will be bound in the tube wall, and it can be almost considered that there is no magnetic field outside the surface. But when there is a defect in the tube wall that cuts the magnetic field lines, because the magnetic permeability of other media at the defect is very small and the magnetic resistance is large, the path of the magnetic field lines will change, which manifests as the leakage of magnetic flux, forming the so-called leakage magnetic field. In actual detection, the magnetization structure carried by the detection equipment is used to magnetize the local tube wall through which the detector passes. When there is a defect in the pipe wall, a leakage magnetic field will be generated, and the probe with a magnetic sensitive sensor located between the two magnetic poles can detect the leakage signal reflecting the defect.

B. Eddy Current Detection

EC detection has been widely used in pipeline detection. It has high detection sensitivity and resolution for pipe wall surface defects. It does not require direct contact with the measured object or couplant during detection. The detection speed is fast and easy to realize automation[7].

EC detection technology is based on the principle of electromagnetic induction. When the detection coil with alternating current is gradually approached to the ferromagnetic tube wall, the magnetic field generated by the coil will cause eddy currents in the tube wall. The size and phase of the eddy current are affected by the properties of the tube wall material and the presence or absence of defects, and a magnetic field will also be generated to react to the magnetic field of the coil to change the impedance of the detection coil. Therefore, it is possible to determine whether the tube wall has defects and its characteristics by detecting the change of the coil impedance [8].

III. Finite Element Simulation Of MFL Detection Of Pipeline Defect

With the rapid development of computer technology, numerical methods are increasingly used in theoretical research of various technologies. The use of engineering software based on numerical models for simulation can improve work efficiency, guide engineering practice, and save experimental costs. In this paper, the static electromagnetic field simulation module of ANSYS software is used to analyze the approximate low-speed MFL detection process [9]. After establishing the pipeline magnetic flux leakage detection entity model, the finite element analysis of the leakage magnetic field at the defect is carried out to explore the influence of the size of the defect on the MFL signal, and to provide theoretical guidance for MFL detection test of the pipeline defects.

A. Simulation Model of MFL Detection of Pipeline Defect

Due to the symmetry of the tested pipeline and the MFL detection device, the magnetic circuit of any part of the pipeline is roughly the same during detection, and 1/4 of the pipeline can be used for simulation, which reduces the scale of modeling and reduces the amount of calculation in the entire simulation process [10].

The created three-dimensional model of MFL detection includes magnets, armature, s, tested pipe (including pipe wall defect), and air fields, as shown in Fig. 1. The model uses magnets to provide the magnetization field for the entire model. The size of the magnet model is $65*40*8mm^3$, and the two magnets have opposite magnetic poles. Their materials are NdFeB with a coercivity of 900kA/m and a remanence of 1.2T. The size of the armature model and the pole piece model are $210*65*30mm^3$ and $65*40*12mm^3$ respectively. Their material are both soft iron, and the property is defined by the B-H curve. The tested pipe has an inner diameter of 377mm and a wall thickness of 8mm. Its material is steel and the property is also defined by the B-H curve.
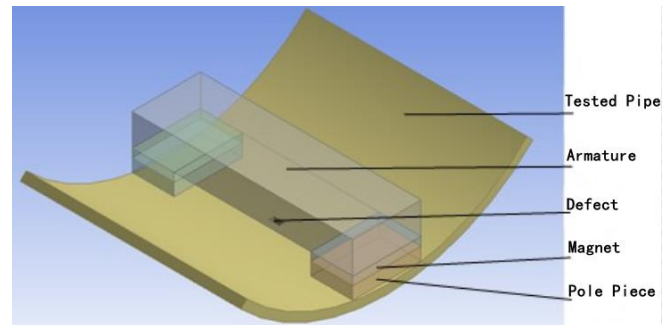
Fig. 1. Simulation model of pipeline defect magnetic flux leakage detection (omit the peripheral air domain)

B. Characteristics of Leakage Magnetic Field at Different Size Defects

In the MFL detection of pipeline defects, the MFL signal waveforms at different types of defects are different. When the size of the same type of defects is different, the leakage magnetic field generated is also different. There are differences in parameters such as the span and amplitude of the signal waveform. This paper sets square defects in the simulation model. The defects of different lengths (along the pipe axis) and different depths (along the pipe radial) are simulated, and the law of the leakage magnetic field at different size defects is analyzed. In order to simulate the travel path and recording waveform of the magnetic sensitive element in the actual detection process, a straight path along the pipe axis is set 1.5mm above the defect, and the magnetic flux density value on the path is extracted.

1) The Influence of Defect Length on Leakage Magnetic Field Signal

In this group of simulations, the length of the defect is set as a single variable, so the widths of the defects are fixed to 5mm, and the depths of the defects are fixed to 8mm, that is, through the pipe wall. The lengths of the defects are respectively taken as 5mm, 8mm, 10mm, 15mm, and 20mm to analyze the change of the leakage magnetic field signal at the defect when the length of the defect changes. The

distribution curves of the axial and radial components of the leakage magnetic field signal (magnetic flux density) at defects of different lengths are shown in Fig. 2.

The distribution curve of the axial component is axisymmetrically distributed at the defect. When the defect length is small, the curve presents a single peak. And when the defect length increases to a certain extent, the curve presents a double peak, that is, a depression appears in the middle. The distribution curve of the radial component is centrosymmetrically distributed at the defect, with positive and negative peaks, and the absolute values of the two peaks are equal.

From the perspective of the curve change law, the length of the defect has an obvious influence on the peak value and span of the axial component. As the length of the defect increases, the peak value decreases and the span increases. The influence law on the peak-to-peak value and the peak-to-peak distance of the radial component is obvious. As the length of the defect increases, the values of both increase.
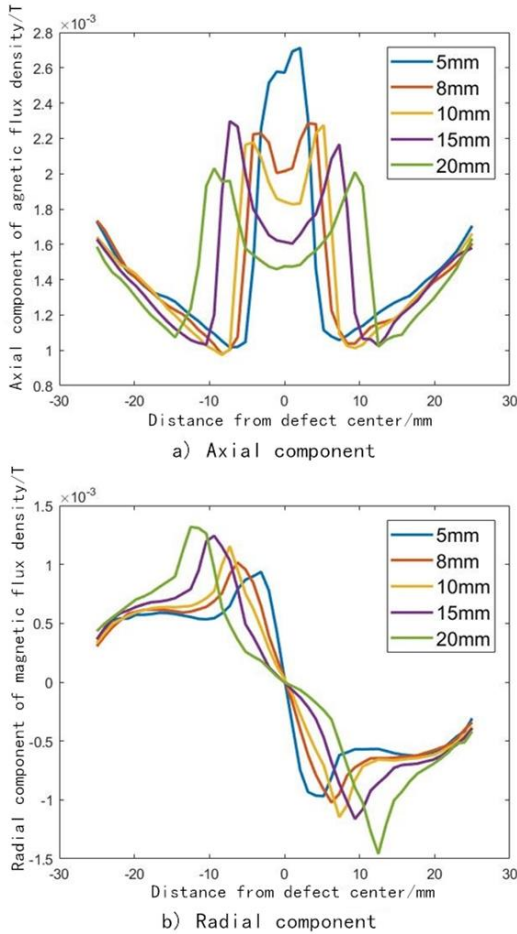
Fig. 2. Distribution curve of magnetic flux density at defects of different lengths (simulation)

2) The Influence of Defect Length on Leakage Magnetic Field Signal

In this group of finite element simulations, the depth of the defect is set as a single variable, so the lengths and widths of the defects are fixed to 5mm and 10mm, respectively. The depths of the defects are respectively taken as 25%, 50%, 75% and 100% (that is, through the tube wall) of the tube wall thickness to simulate and analyze the influence of the defect depth on the leakage magnetic field at the defect. The

distribution curves of the axial and radial components of the leakage magnetic field signal (magnetic flux density) at defects of different depths are shown in Fig. 3.

From the change law of the curve, the influence of the defect depth on the peak value of the axial component is obvious. As the depth of the defect increases, the peak value increases. It also affects the peak-to-peak value and the peak-to-peak distance of the radial component. As the depth of the defect increases, the peak-to-peak value increases, and the peak-to-peak distance decreases.
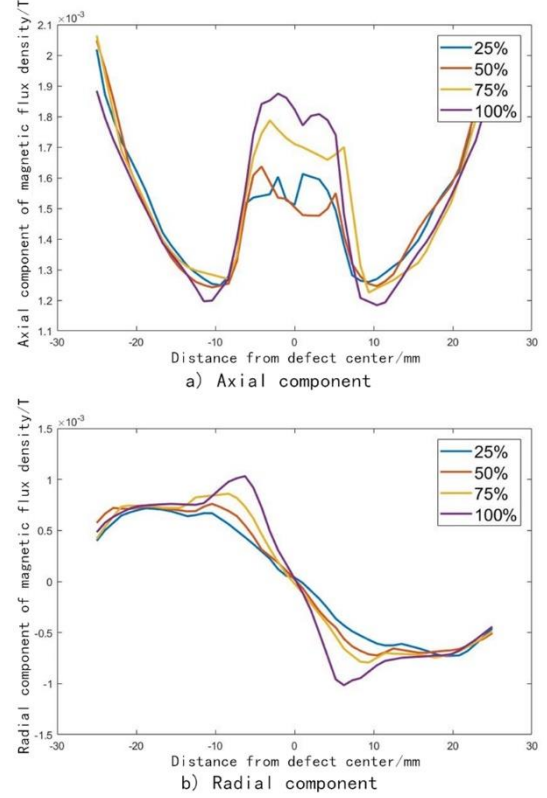
Fig. 3. Distribution curve of magnetic flux density at defects of different depths (simulation)

IV. DESIGN AND REALIZATION OF IN-PIPE DETECTION SYSTEM

The in-pipe detection system designed and implemented in this paper mainly includes detection sensors and data acquisition and storage system, which can rely on conventional pipeline detection equipment. Detection sensors mainly include MFL sensors and EC sensors, as well as auxiliary sensors such as attitude sensors and temperature sensors. The data acquisition and storage system is responsible for real-time acquisition, processing and storage of sensor detection data. The overall structure of the in-pipe detection system is shown in Fig. 5.

A. Detection Sensors

The detection system in this paper uses three-axis high-definition MFL detection technology and single-frequency EC detection technology as the main detection methods. Therefore, the MFL detection sensor and EC detection sensor are the core parts of the system. In order to realize the integration and miniaturization of the detection probe, the two sensors are integrated on the same detection probe. There are four three-axis MFL detection sensors and two EC detection

sensors on one probe. The specific structure design of the detection probe is shown in Fig. 4.
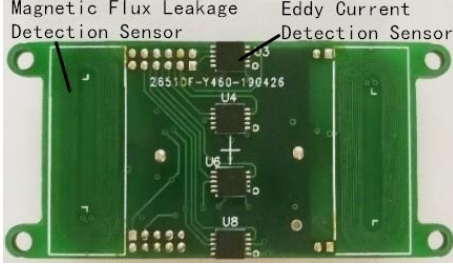
Fig. 4.   MFL and EC detection probe

The MFL detection method is sensitive to volumetric defects, but it is not sensitive to parallel cracks, nor can it distinguish between inner and outer wall defects. The EC detection method is sensitive to near-surface defects such as cracks, but is greatly affected by lift-off. Therefore, the design of the above-mentioned integrated probe is beneficial for the two detection methods to complement each other, and better complete the pipeline detection task [11]. First of all, the two can cooperate with each other to distinguish the inner and outer wall defects. When the two detection methods measure the tube wall defects at the same time, the defect is the inner wall defect, otherwise it is the outer wall defect. Secondly, the EC detection method can make up for the lack of crack detection function of the MFL detection method. Finally, the EC detection data can correct and compensate the MFL detection data, and the lift-off value contained in the EC detection data can be used to improve the quantification accuracy of the defect.

In addition, the attitude sensor can be used to record the attitude changes of the internal detection equipment during the detection process, especially the roll angle needs to be used for defect location. The temperature sensor can be used to record changes of the ambient temperature during the detection process.

B. Data Acquisition and Storage System

In the in-pipe detection system, the number of MFL and EC detection probes should be designed according to the size of the pipeline to be tested, and the probe or sensor should be covered along the circumference of the pipeline as far as possible. In practical detection, more sensors and higher sampling rate mean higher detection accuracy. So a large amount of data will be produced in the detection process. The real-time and accurate acquisition and storage of sensor data is the premise and foundation of pipeline defect analysis. In this paper, a special data acquisition and storage system is designed and implemented.

The system includes data acquisition module, data summary module and data storage module. The detection data from sensors are collected, summarized and stored by three modules in turn. The core of the data acquisition module is STM32 microcontroller, which is integrated in the MFL and EC detection probe to collect the data of the MFL and EC sensors in real time. The core of data collection module is FPGA. Considering the large number of probes, it is designed as a hub with two or more stages. In this module, the detection data are collected into the FPGA of the final hub layer by layer, and then merged into a complete frame, and the frame head, frame tail and frame number are added. In addition, the data of attitude sensor and temperature sensor are directly collected by FPGA on the final hub. The data storage module takes multi-core ARM processor as the core, runs Linux system and mounts SSD. In this module, the data frame is further packaged and stored in the solid-state disk. The data stored in the hard disk can be transmitted to the host computer through the network port for further processing and analysis. In order to ensure the real-time storage of a large number of detection data, two or more SPI buses can be selected to improve the transmission speed.
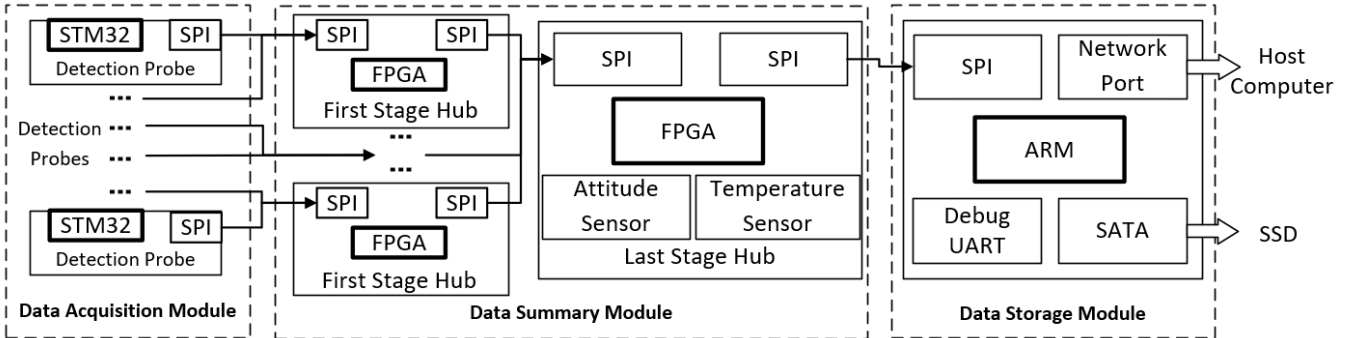
Fig. 5.   Overall structure of in-pipe detection system

V. Verification And Analysis Of Pull Test

In order to verify the performance of the in-pipe detection system in this paper, the in-pipe detection equipment is used to carry out the pull test in the special pipeline, and the detection data are analyzed. The inner diameter of the test pipe is 508 mm and the wall thickness is 7.1 mm. There are girth weld defects, spiral weld bead defects, metal loss defects, oil stealing hole and inner and outer wall defects. In this paper, the length of the defect is along the axial direction, the width is along the circumferential direction, and the depth is along the radial direction.

A. Detection of Metal Loss Defects

There are 32 rectangular metal loss defects with different length, width and depth on the test pipeline. The MFL detection sensor can effectively detect these defects. The signals of MFL detection of metal loss defects show that defects with depth of 40% or above, length of 5mm or above, or width of 20mm or above can be effectively detected. In other words, only when the depth of the defect is shallow and the length and width are small, the MFL detection cannot detect it. The MFL detection signals at one group of defects are shown in Fig. 6. The depths of the defects in this group are 60%, and the lengths and widths are different.

Similar to the above finite element simulation analysis, two groups of MFL detection signals at defects are selected for comparative analysis. For one group of defects, the lengths are 5 mm, 20 mm and 40 mm, the widths are 20 mm and the depths are 60%. For the other group, the depths are 20%, 40%, 60% and 80% respectively, and the lengths are 40 mm and the widths are 40 mm. The distribution curves of axial and radial components of detection signals at defects with different lengths are shown in Fig. 7. The distribution curves of axial and radial components of detection signals at defects with different depths are shown in Fig. 8. The analysis shows that the MFL detection signal at the defects has obvious characteristics, and has obvious change law with the change of defect length and depth, which is basically consistent with the above results of finite element simulation. In addition, the number of sensors detected defects is positively correlated with the width of the defect.
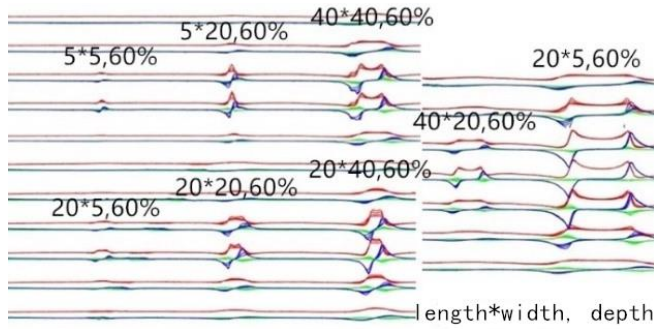
Fig. 6. MFL detection signals at metal loss defects
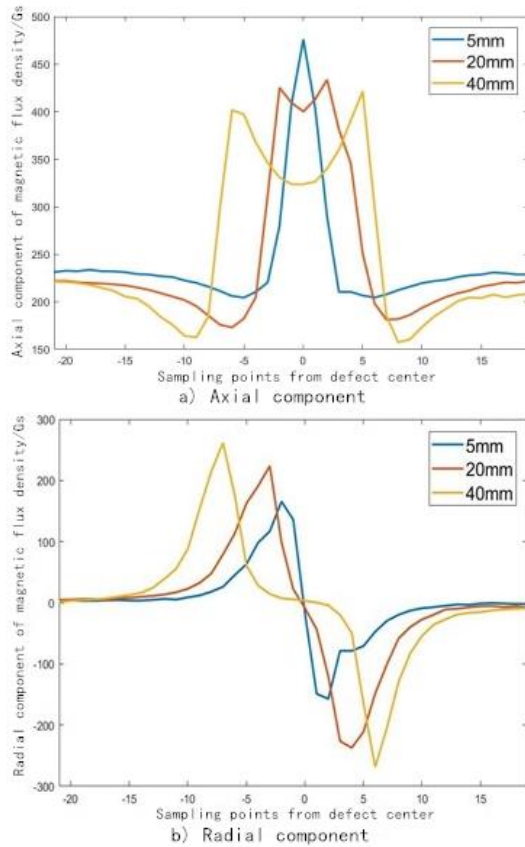
Fig. 7. Distribution curve of magnetic flux density at defects of different lengths (actual measurement)
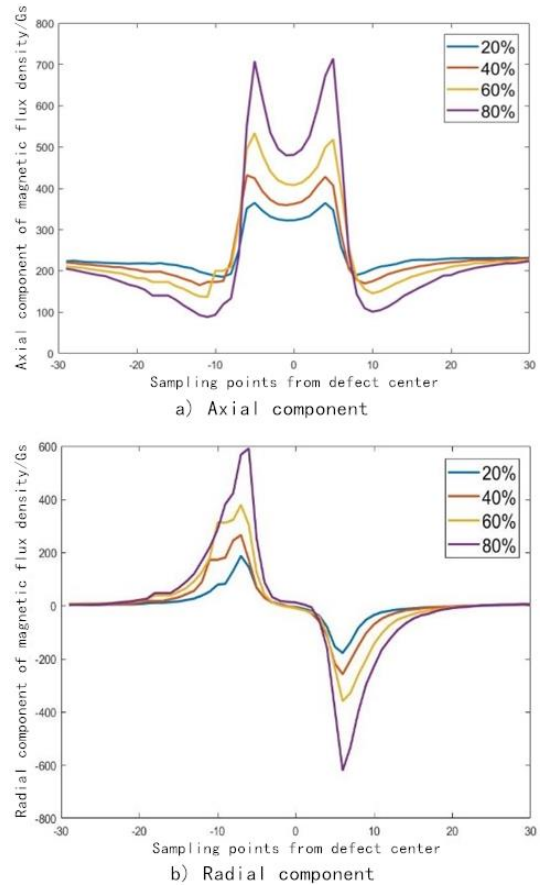
Fig. 8. Distribution curve of magnetic flux density at defects of different depths (actual measurement)

B. Detection of Inner and Outer Wall Defects

There are inner and outer wall defects on the test pipe, which can be distinguished by the cooperation of MFL and EC detection sensors. The MFL and EC detection signals at the inner and outer wall defects are shown in Fig. 9. The difference between the MFL detection signals at the inner and outer wall defects is not obvious, while the EC signals are significantly different. In other words, the EC detection method is only sensitive to the inner wall defects, but the EC signals cannot reflect the details of the defects.
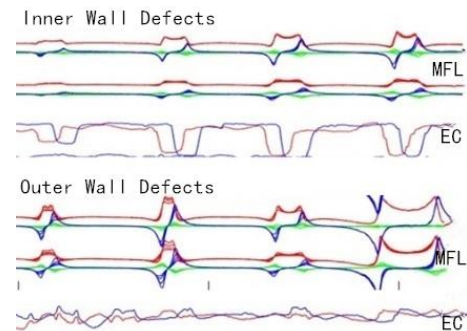
Fig. 9. MFL and EC detection signals at inner and outer wall defects

C. Detection of Girth Weld Defects

There are 20 girth weld defects with different lengths, widths and depths on the test pipeline. The MFL detection sensor can effectively detect such defects. The MFL detection signals of girth weld defects show that the defects with depth of 40% and above can be effectively detected, and the

minimum length of the defects that can be detected is 0.5mm. Furthermore, the MFL signal changes with the lengths and depths of defects, and the number of sensors detected defects is positively correlated with the defect width. The MFL detection signal of one group of defects is shown in Fig. 10. The length of the defects are 5mm, the widths are 200mm, and the depths from left to right are 20%, 40%, 60%, 80% and 100%.
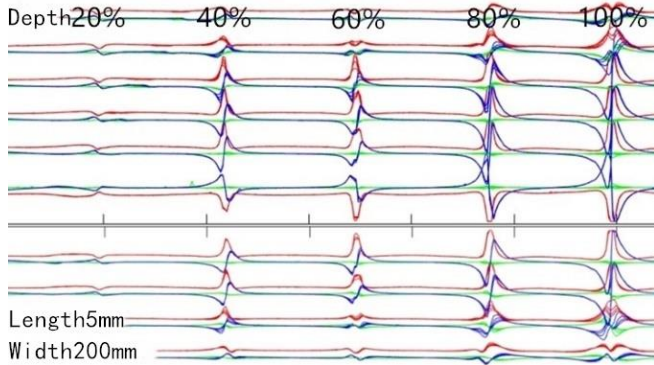
Fig. 10. MFL detection signals at girth weld defects

VI. Concluding Remarks

In this paper, an in-pipe detection system based on MFL and EC detection is developed. The detection probe integrated with MFL and EC sensors and the corresponding data acquisition and storage system are designed and implemented.

By comparing and analyzing the results of the finite element simulation and the pull test of pipeline defects, the characteristics and variation rules of MFL signals at different sizes of defects are obtained. The peak value, span of the axial component and the peak to peak distance of the radial component of the signal can reflect the defect size.

The functions of the system are verified by the pull test in the pipeline with artificial defects. The detection sensors and data acquisition and storage system meet the design requirements.

The system can effectively detect girth weld defects, spiral weld bead defects, metal loss defects, drilling and oil stealing hole, and can distinguish the inner and outer wall defects.

REFERENCES

[1] Kandroodi M R, Araabi B N, Ahmadabadi M N, Shirani F, Bassiri M M. Detection of natural gas pipeline defects using magnetic flux leakage measurements[C]// Electrical Engineering. IEEE, 2013:1-6.

[2] Li Xiang, Gao Anjie, Cheng Jian, Zhang Chunhe, Ma Weiping. Research on safety management system of long distance pipeline in north america[J]. Automation in Petro-Chemical Industry, 2017, 53(5):1-4,14.

[3] Yang Lijian, Geng Hao, Gao Songwei. Magnetic flux leakage internal detection technology of the long distance oil pipeline[J]. Chinese Journal of Scientific Instrument, 2016, 37(8):1736-1746.

[4] Kong Chaojin, Hu Lifeng, Fu Zhenlin, Niu Zhiyong, Yu Fangyong. The discussion on technology status and development trend of pipeline internal inspection[J]. Total Corrosion Control, 2019, 33(8):4-8.

[5] Bao Qingjun, Shuai Jian. Research progress in the inner detection technology of oil and gas pipeline[J]. Contemporary Chemical Industry, 2017, 46(2):298-301.

[6] Liu Bin, Yang Lijian. Magnetic flux leakage internal detection technology of the long distance oil pipeline[M]. Beijing: China Machine Press, 2017:34.

[7] Li Xingang. NDT technology and equipment analysis of oil and gas pipelines[J]. Petrochemical Industry Technology, 2017, 24(4):94-95.

[8] Huang Songlin. Theory and application of magnetic flux leakage internal detection for oil and gas pipeline defects[M]. Beijing: China Machine Press, 2013:9.

[9] Peng Jiageng, Qing Siji, Kong Xiaohua, Lin Qingli, Chengfenglan. Analysis of permanent-electromagnet clamping technique and its influencing factors[J]. Journal of Plasticity Engineering, 2011, 18(2):18-23.

[10] Zheng Keyao. Detection technology in magnetic flux leakage method on the pipeline of 300mm in diameter[D].

[11] Wang Shaoping, Wang Zengguo, Liu Jinhai, Su Hanguang, Wang Gang. Identification method between internal defect and external defect for pipeline based on tri-axial magnetic flux leakage and electric eddy current inspection[J]. Control Engineering of China, 2014, 21(4):572-578.

Fault Accurate Location Method in UHV GIL Based on Transient Voltage Travelling Wave

Yong Wang
Guangzhou Power Supply Co. Ltd
Electric Power Test & Research Institute
Guangzhou, China
wangy@guangzhou.csg.cn

Dengwei Ding
Sichuan Energy Internet Research Institute
Tsinghua University
Chengdu, China
sunnyall123@163.com

Lu Zhu
Guangzhou Power Supply Co. Ltd
Electric Power Test & Research Institute
Guangzhou, China
zhulu19870408@126.com

Wenxiong Mo
Guangzhou Power Supply Co. Ltd
Electric Power Test & Research Institute
Guangzhou, China
gzmwx@139.com

Weinan Fan
Guangzhou Power Supply Co. Ltd
Electric Power Test & Research Institute
Guangzhou, China
fwn_ee@foxmail.com

Ziwei Zhang*
Sichuan Energy Internet Research Institute
Tsinghua University
Chengdu, China
ziwei.z@outlook.com

Abstract—In case of insulation breakdown accident of ultra-high voltage (UHV) gas insulated metal-enclosed transmission line (GIL), accurate location of breakdown position is conducive to emergency repair and rapid recovery of power transmission. This paper proposed the fault location method based on transient voltage travelling wave, which is actual applied in 1100kV Sutong GIL utility tunnel project. During the operation, when the GIL flashover occurred , the transient voltage generated during the breakdown was completely recorded. Through waveform analysis, it is found that when UHV GIL insulation breaks down, the first change steepness of transient voltage is less than 700ns, and spreads between the break point and bushing, form a typical traveling wave process. Through the time when the transient traveling wave first reached the measuring points at both ends of GIL and combined with the propagation speed of traveling wave, the accurate position of the fault point is calculated and the fault point is quickly disassembled for maintenance. According to the results of disintegration verification, the positioning error of the system is less than 5 meters. It can be seen that the adoption of wide-band transient voltage sensing can realize the rapid and accurate location of insulation breakdown fault of UHV GIL, which is conducive to improving the operation reliability of UHV transmission system.

Keywords—GIL; insulation breakdown; transient voltage travelling wave; accurate location

I. Introduction

Ultra-high voltage (UHV) gas insulated metal-enclosed transmission line (GIL) is an important part of the transmission system, its reliable operation is directly related to the security and stability of the power system [1]. After the insulation breakdown of UHV GIL, in order to reduce the impact of power failure, it is necessary to locate the discharge location quickly and accurately. Due to the large number of gas chambers in long-distance GIL, it is not practical to directly use the conventional gas decomposition components detection method to locate the fault chamber. Therefore, it is essential to study the fast and accurate positioning method for UHV GIL equipment [2]-[4]. At present, the research on fault location technology of transmission lines mostly focuses on the conventional overhead lines and cables, and the location accuracy is mostly about 500m, which is difficult to meet the

requirements of the fault location accuracy in UHV GIL [5]-[8]. If other GIS equipment fault location technologies are applied to UHV GIL. In order to achieve 20 meters positioning accuracy, hundreds of sensors need to be arranged. The reliability of the monitoring system is difficult to guarantee, which not only makes the project cost unacceptable, but also brings huge operation and maintenance burden.

Through the real full-scale test of 1100kV GIL equipment, it was found that when the insulation breakdown occurs inside the equipment, the transient voltage with the sharp change time of wave head less than 150ns will be generated. In this paper, a GIL fault accurate location method based on transient voltage travelling wave is proposed, and it is applied in 1100kV Sutong GIL utility tunnel project.

The transient voltage travelling wave generated when insulation breakdown occurs is monitored. The time difference and wave characteristics of transient voltage waveform arriving at measuring point in the GIL are accurately obtained by using high-precision GPS module, and the exact location of the fault is calculated by combining the propagation speed of voltage traveling wave. During the operation of GIL, the equipment breaks down. The fault accurate location system accurately triggers and automatically locates, which verifies that all performance and positioning errors of the system meet the requirements of field application.

II. Location Theory Based on Travelling Wave

During the operation, because the insulation medium of 1100kV GIL is usually SF_6 gas and the insulation distance is short, the insulation breakdown inside the equipment is different from the breakdown flashover on the conventional overhead line, which will result in more high-frequency components of electromagnetic wave produced by breakdown discharge in GIL and steeper wave head. By building an 1100kV GIL full-scale test platform to simulate the surface flashover of spacer, the measured transient process of single breakdown is shown in Figure 1. It can be seen that the rising edge of transient voltage of single breakdown is about 110ns, which makes it possible to improve the fault location accuracy of GIL.
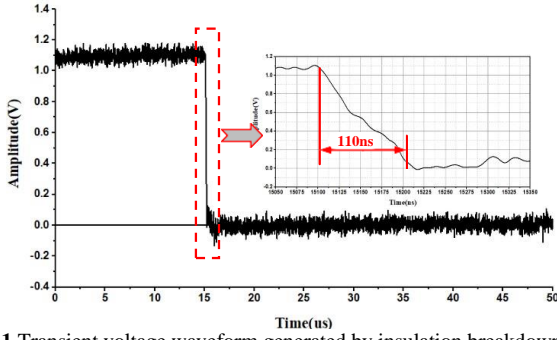
Fig. 1 Transient voltage waveform generated by insulation breakdown of 1100kV GIL.

A. Two-terminal Location Method

The principle of two-terminal location method based on transient voltage traveling wave in GIL is shown in Figure 2. It can be seen in the figure that when the voltage is u at the time of T, the insulation breakdown will produce the same amplitude as the breakdown voltage, and the transient voltage traveling wave with opposite polarity will spread rapidly to the north and south ends of GIL. The distance between the breakdown point and the north and south sensors is L_{SS} and L_{SN} respectively. Combined with the high-precision timing module, the time when the transient voltage waveform first reaches the sensors can be determined as t_s and t_n respectively. According to the propagation speed V of the transient voltage traveling wave in GIL and the distance L between the south and north sensors, the L_{SS} can be calculated according to formula (1).

$$L_{SS} = \frac{(t_s - t_n) \times v + L}{2} \tag{1}$$
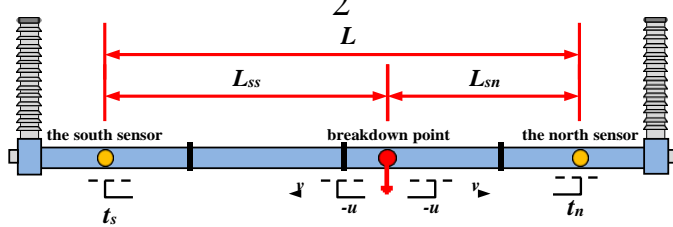
Fig. 2 Schematic diagram of two-terminal traveling wave location principle of GIL

B. Single-terminal loaction method
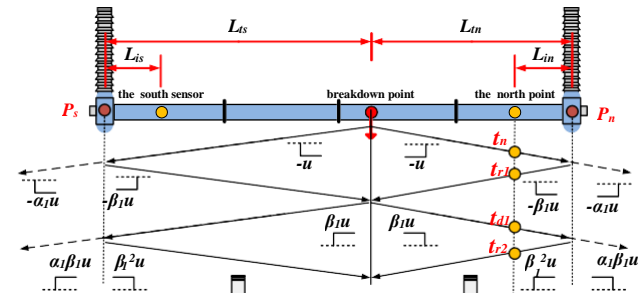
Fig. 3 Schematic diagram of single-terminal traveling wave location principle of GIL

The transient voltage generated during the internal insulation breakdown propagates in GIL as shown in Figure 3. In the figure, the end of the north and south bushing is represented as P_s and P_n respectively, the distance between the breakdown point and the end of the north and south bushing is L_{ts} and L_{tn} respectively, and the distance between the measuring points and the end of the bushing is L_{is} and L_{in}.

respectively. Because the propagation law of the transient voltage waveform is basically the same in the north and south sides, this paper only analyzes the propagation process of the transient voltage in the north side, and ignores the attenuation caused by the dielectric loss in the propagation process. It can be seen in the figure that the transient voltage with an amplitude of $-u$ at t_n firstly propagates from the fault point to the north measuring point, and then continues to propagate to the end of the bushing.

As the end of the bushing is connected with the overhead line, the wave impedance changes abruptly, and the transient voltage will reflect at P_n. The equivalent circuit is shown in Figure 4. The reflection coefficient α_1 and β_1 are calculated according to formula (2). Where Z_G represents wave impedance of the GIL, usually tens of ohms, Z_L represents wave impedance of overhead line, usually hundreds of ohms, Z_C represents wave impedance of GIL enclosure to the ground, usually thousands of ohms. The reflected transient voltage with an amplitude of $-\beta_1 u$ propagates back from P_s to the north measuring point at t_{r1}, and then to the fault point again. At this time, the fault point is short circuited to the ground, as shown in the formula (3), the refraction coefficient α_2 is 0, the reflection coefficient β_2 is -1, and the waveform is fully negative reflection. The reflected transient voltage with the amplitude of $\beta_1 u$ propagated to P_n again, and reached the north measuring point at t_{d1}. Then the reflection occurs again at P_n, and the reflection transient voltage with the amplitude of $\beta_1^2 u$ reaches the north side for measurement at the moment t_{r2}.

1. bushing port　　　**2. breakdown point**
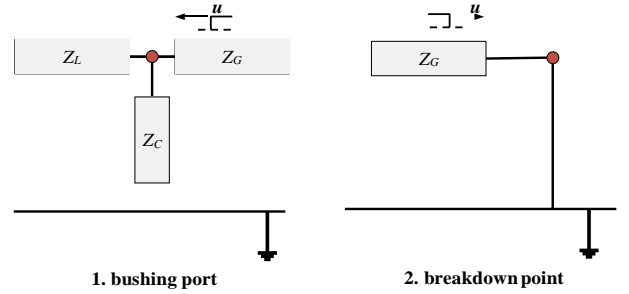
Fig. 4 Equivalent circuit diagram of bushing terminal and breakdown point
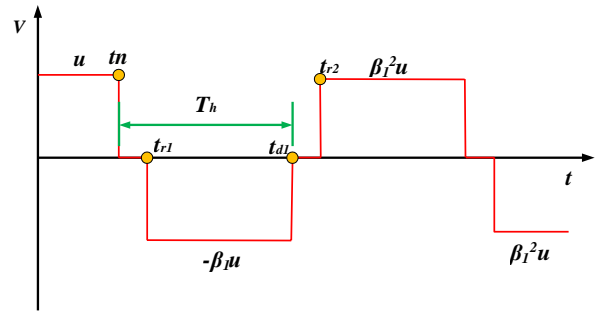
Fig. 5 Theoretical waveform of transient voltage at the south measuring point

$$\alpha_1 = \frac{2 Z_L Z_C}{Z_G Z_L + Z_G Z_C + Z_L Z_C}$$

$$\beta_1 = \frac{Z_L Z_C - Z_G Z_L - Z_G Z_C}{Z_G Z_L + Z_G Z_C + Z_L Z_C} \tag{2}$$

$$\alpha_2 = \frac{2 \times 0}{0 + Z_G} = 0; \quad \beta_2 = \frac{0 - Z_G}{0 + Z_G} = -1 \tag{3}$$

According to the transient voltage propagation process in Figure 3, the theoretical waveform of the transient voltage

at the north sensor can be obtained, as shown in Figure 5. It can be seen in the figure that the transient voltage propagation distance from t_n to t_{d1} is twice the distance from the breakdown point to the end of the north bushing. In this paper, the period of square half wave (T_h) is defined $T_h=t_{d1}-t_n$. Therefore, the distance L_{tn} from the breakdown point to the end of the north bushing can be calculated according to formula (4), so as to realize the single terminal location of the fault point.

$$L_{tn} = \frac{T_h \times v}{2} \qquad (4)$$

III. APPLICATION OF ACCRUATE LOCATION METHOD
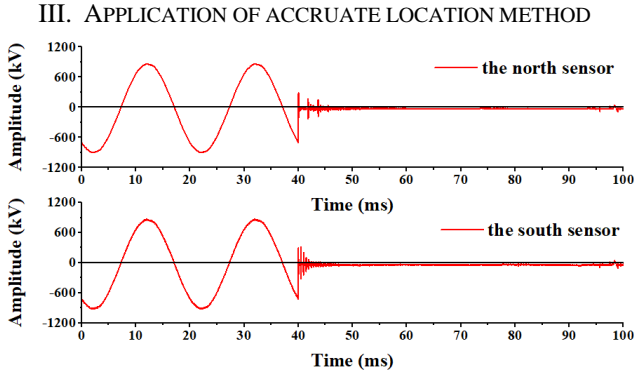
Fig. 6 Fault voltage waveform recorded by transient voltage system during flashover of 1100kV GIL

During the actual normal operation, the internal insulation of GIL breaks down. The fault voltage waveform within 100ms recorded by the transient voltage monitoring system during the flashover of the spacer in operation is shown in Figure 6. In the actual operation, when the bushing is connected with the overhead line, the transient voltage propagates from the fault point to the bushing, and the components reflected back to the fault point form a typical square wave voltage traveling wave. The transient components refracted to the overhead line will be transmitted to the substation, reflected again, and then back to GIL, refracted to the measuring point through the bushing. Therefore, when the spacer flashover in operation, the voltage rapidly drops to zero after a short period of oscillation, and then there will be several transient voltages reflected back from the substation. It can be seen that since the south end of GIL is about 60km away from the substation, five obvious transient voltage waveforms are detected at the south side of the substation, with a duration of about 3 ms. The north end of GIL is about 260 km away from the substation. Three obvious reflected transient voltage waveforms have been detected at the north measuring point, with a duration of about 5 ms.

The first transient voltage recorded at the north and south measuring points is expanded as shown in Figure 7. It can be seen that the transient voltage recorded at the south bank measurement point lasts approximately 80 us, and there are about three obvious square wave changes. The transient voltage recorded at the north bank observation point lasts approximately 150 us, and there are about five obvious square wave changes. The duration of transient voltage is obviously shorter than that of insulation breakdown. The initial square half wave period of the voltage measured at the south measuring point is 12.14 us, the voltage wave head rises from -710 kV to 0 kV, and the half wave steep change

time is about 640 ns. The initial square half wave period of the voltage measured at the north measuring point is 26.83 us, the voltage wave head rises from -710 kV to 0 kV, and the half wave steep change time is about 760 ns. The abrupt change of transient voltage waveform is slower than that of insulation breakdown.
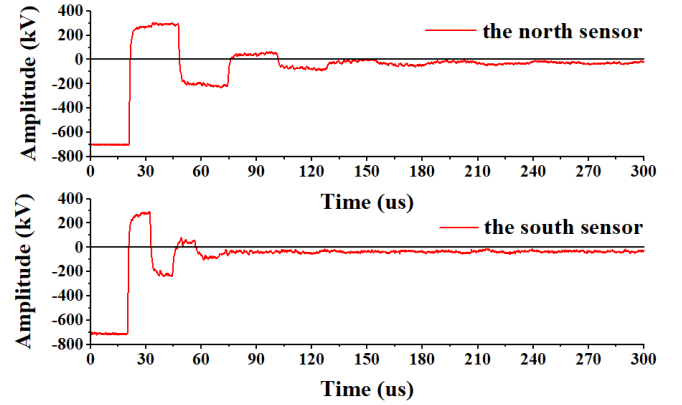
Fig. 7 Complete transient voltage waveform induced by GIL first breakdown flashover of 1100kV GIL
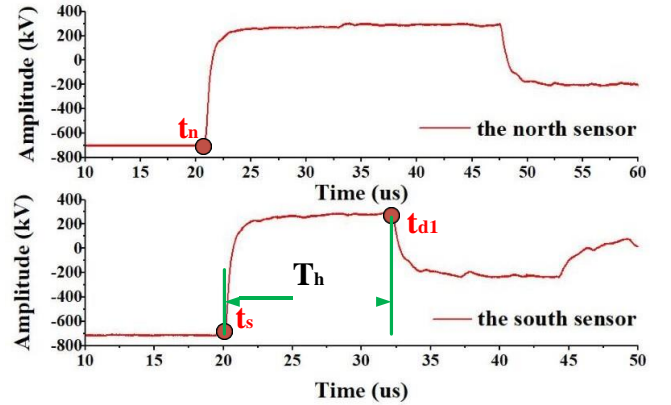
Fig. 8 Front expansion of transient voltage waveform

From Figure 8, the relative position between the breakdown transient voltage t_{d1} and t_s can be determined, and the square half wave period th is about 12.14 us. According to the equivalent capacitance (45 pF/m), equivalent inductance (0.26 uH/m) and equivalent resistance (3.33 $u\Omega$/m) provided by GIL manufacturer, the theoretical calculation shows that the propagation speed V of transient voltage in the 1100kV GIL is 292.4 m/us. The propagation velocity V is further corrected to 292.553 m/us by using the results of previous field positioning. The distance between the north and south measuring points of the GIL is 5657.7m, and the distance between the sensor and the end of the casing is 8.36m. Therefore, according to the single end location algorithm, the distance from the fault point to the south end of the bushing L_{ts} is calculated as follows:

$$L_{ts} = \frac{T_h \times v}{2} = 1775.8m$$

The point where the transient voltage changes obviously for the first time is selected as the starting point of the transient voltage, and the corresponding time is the time when the traveling wave reaches the measuring point. As shown in Figure 9, according to the data recording information of the online monitoring system, the accurate time t_s of the voltage traveling wave reaching the south sensor is 11:48:26:388561626 ns. the accurate time t_n of the voltage traveling wave arriving at the north sensor is

11:48:26:388568876 ns. According to the two terminal positioning algorithm, the distance from the fault point to the south end of the casing is calculated as follows:

$$L_{ts} = L_{is} + L_{ss} = L_{is} + \frac{(t_s - t_n) \times v}{2} = 1776m$$
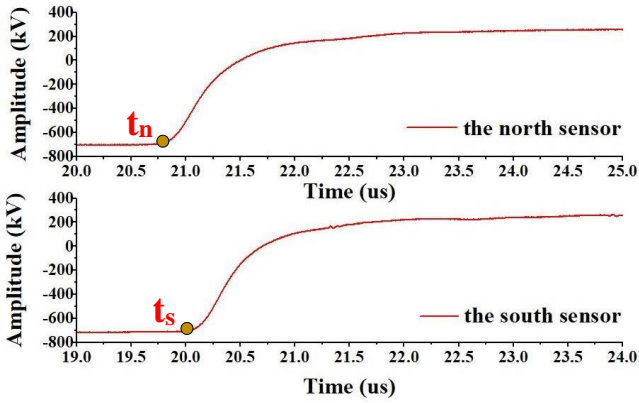
Fig. 9 Rising edge and recording trigger point of transient voltage waveform
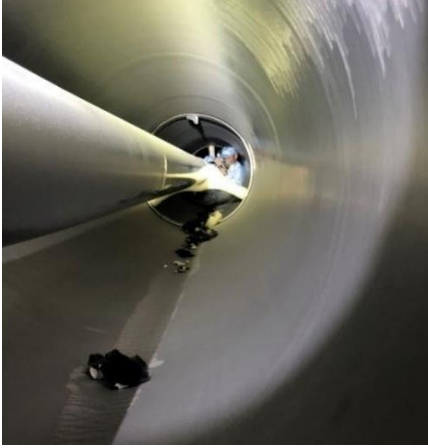
Fig. 10 Disintegration verification of GIL failure point

It can be seen that the fault location obtained by single-terminal and two-terminal location algorithm is basically the same, with a difference of 0.2m. According to the positioning results, the GIL equipment was serviced, and the breakdown point was found at the support insulator, 1772 meters away from the end of the south bushing. The three post insulator discharges the GIL enclosure, and the insulator has cracked, as shown in Figure 1. The difference between the actual fault location and the result of single-terminal algorithm is about 3.8m, and the difference between the results of two-terminal algorithm is about 4m. It can be seen that the single-terminal and two-terminal location algorithm can realize the fault accurate location of UHV GIL, and the positioning error is slightly less than 5m, which meets the engineering requirements.

IV. CONCLUSION

In this paper, the propagation characteristic of transient voltage traveling wave generated by insulation breakdownin UHV GIL is deeply analyzed, and the single-terminal and two-terminal location algorithms are proposed to realize the fault accurate location in GIL, and a precise location system based on transient traveling wave is constructed, which is applied to 1100kV Sutong GIL utility tunnel project. When the insulation breakdown occurs in the actual operation of GIL, the system can locate accurately. It is found that the duration of single transient voltage caused by flashover of post insulator is short, about 150 us, and the first half wave steep change time is about 760 ns. The results show that the positioning errors of single-terminal and two-terminal location algorithms are almost equal, both less than 4 meters. It can be seen that the proposed method of UHV GIL accurate location based on transient voltage traveling wave monitoring can effectively and accurately monitor the transient voltage generated by fault breakdown in UHV GIL, and achieve the precise positioning with error less than 5 meters, which can effectively reduce the troubleshooting time of UHV GIL shorten rush repair period, and improve the operation reliability of UHV transmission system.

REFERENCE

[1] QI Bo, ZHANG Guixin and LI Chengrong, et al. Research Status and Prospect of Gas-insulated Metal Enclosed Transmission Line [J]. High Voltage Engineering. 2015，41(5)：1466-1473.

[2] CHEN Xuanshu, HU Yi, XIN Yaozhong, et al．Prospect of high voltage long distance compressed-air insulated transmission lines [J]．High Voltage Engineering, 2009, 35(12)：3137-3142.

[3] LU Jiaming, CAO Weiwei, ZHOU Zhenhua. The structural design of 550 kV GIL bus [J]. Electrical Engineering, 2015, 16(9)：59-63.

[4] CHEN Zheng , SU Jin-xi and WU Xin-rong, et al. Fault Location Algorithm for High Voltage Transmission Line Based on Distributed Parameter [J]. Power System Technology, 2000 , 24(11)：31～32.

[5] XIA Lu-lu, HE Zheng-you and ZHANG Jun. Adaptability Analysis on Fault Location of Continuous Railway Power Line by Traveling Wave Front Extraction Algorithm Based on Mathematical Morphology [J]. Power System Technology, 2009, 33(8)：78-83.

[6] CHUL-HWAN Kim, HYUN KIM and YOUNG-HUN KO, et al. A Novel Fault-Detection Technique of High-Impedance Arcing Faults in Transmission Lines Using the Wavelet Transform [J]. IEEE Transactions on Power Delivery, 2002, 17(4): 921-928.

[7] LI Yang , HUANG Ying and CHENG Lexiang. Fault locating with coordination of travelling wave arrival time and velocity [J]. Electric Power Automation Equipment, 2010, 30(11)：44-47.

[8] WU Lin-yong, HE Zheng-you and QIAN Qing-quan. A Single Ended Fault Location Method Using Traveling Wave Natural Frequency. Proceedings of the CSEE, 2008, 28(10)：69-75.

An Optical Partial Discharge Localization Method Based on Simulation and Machine learning in GIL

Yiming Zang
Department of Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
zangyiming@sjtu.edu.cn

Yong Qian*
Department of Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
qian_yong@sjtu.edu.cn

Hui Wang
Department of Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
wanghui8203@sjtu.edu.cn

Antian Xu
Department of Light Source and
Illuminating Engineering
Fudan University
Shanghai, China
xuantian1996@163.com

Gehao Sheng
Department of Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
shenghe@sjtu.edu.cn

Xiuchen Jiang
Department of Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
xcjiang@sjtu.edu.cn

Abstract—As gas-insulated transmission line (GIL) is widely used, the partial discharge (PD) phenomenon that occurs during their operation is one of the main reasons for the deterioration of their insulation status. Therefore, the detection and localization of PD in the GIL plays an important role in ensuring the safe and stable operation of the equipment. At present, the PD localization methods for GIL are mainly ultrahigh frequency (UHF) method and ultrasonic method, while these methods are susceptible to mechanical vibration and electromagnetic interference. Optical detection, as a sensitive and effective PD detection method, is rarely used in the field of GIL PD localization. Accordingly, this paper proposes a GIL PD localization method based on optical PD simulation and machine learning. This method establishes a simulation model that is exactly the same as the actual GIL in terms of structure, size and sensor arrangement, where the PD optical simulation experiment is performed to build a PD simulation fingerprint database. Each fingerprint in the fingerprint database corresponds to a PD source location information. Based on this, the PSO-SVM machine learning algorithm is used to match the actual PD fingerprints with the fingerprints in the simulation fingerprint database to obtain the localization results. This method overcomes the difficulty of collecting a large amount of field data to build a fingerprint database in the existing optical fingerprint localization method through simulation. And the structure of GIL in the simulation can be customized according to the actual equipment, while the workload of obtaining the fingerprint database through actual experiments is very heavy for different types of equipment. Through experimental verification, this paper selects 12 typical locations of PD sources in the GIL experimental tank for testing. The average localization error is 10.58mm.

Keywords—partial discharge, fingerprint localization, gas-insulated transmission lines, optical detection, machine learning

I. INTRODUCTION

Since the 1970s, the GIL has been proposed and invented. Due to the excellent electrical properties, such as low loss and high transmission capacity, it is gradually being widely used in the power transmission [1]. However, in the process of live operation of GIL equipment, partial discharge (PD) that endangers the safe operation of GIL occurs due to the deterioration of insulation or defects of equipment [2].

In order to avert the hazards caused by PD, the timely and effective detection and localization of PD is of great significance to GIL, especially the localization which can assist maintenance personnel in troubleshooting. The current localization methods for PD are mainly based on ultrahigh frequency (UHF) signals, ultrasonic signals and electrical signals [3]. The localization method based on UHF signals mostly adopts the method of solving nonlinear positioning equations based on Time Difference of Arrival (TDOA). However, due to the high speed of UHF signal propagation and susceptibility to electromagnetic interference, the response speed and anti-interference performance of its detection system are relatively high, so it is easy to introduce system errors in actual detection. Moreover, the iterative solution process of nonlinear equations is relatively sensitive and complicated. Both measurement and system errors may cause the equations to have no solution, resulting in incorrect results [4]. The PD localization based on ultrasonic signals mainly uses the array sensor and acoustoelectric joint TDOA positioning method. However, the attenuation rate of ultrasonic waves in space is fast, and the wave speed is easily affected by factors such as air pressure, temperature, mechanical vibration, and medium, which is only suitable for locating in a small range [5]. The PD localization based on electrical signals can only provide the electrical position of the PD, and cannot guide the determination of the location of the fault space during the maintenance process, which is difficult to be promoted and used on site.

PD localization based on optical signals has the advantages of anti-electromagnetic interference, anti-mechanical vibration interference, high sensitivity and fast response speed, which can effectively detect the occurrence of PD in GIL and have broad development prospects and application value [6]. However, there are relatively few studies on PD localization based on optical signals. A fiber array technology is used for PD localization, while the detection range of this method is limited [7]. The machine learning is introduced into PD localization in the paper [8]. In the early stage of localization, a large number of PD experiments are used to obtain the PD fingerprint database containing the location information of the PD source, which is used as the training set of machine learning algorithms. During the location process, the collected optical PD fingerprints are matched with the information in the fingerprint database through the machine learning algorithm to obtain the location of the PD source. This method requires a large number of real PD experiments in the GIL, and it turns out to be difficult to conduct such a large number of PD experiments in actual equipment. Besides, this method can only identify the specific PD sources at locations where PD experiments have been conducted, reducing the localization accuracy.

Therefore, we propose a PD localization method based on optical simulation and machine learning, which introduces the simulation model into the PD localization to replace the actual PD experiment. Firstly, a simulation model that is exactly the same as the actual GIL tank size, sensor location, and material is established. Secondly, through the simulation model, the optical signal strength received by each sensor when PD occurs in any location in the GIL tank is obtained, and an optical simulation fingerprint database is constructed based on this. Finally, the detected optical PD fingerprints are matched with the fingerprints in the optical PD fingerprint database through the machine learning algorithm that is particle swarm optimization-support vector machine (PSO-SVM) used in this paper to obtain the actual PD location.

The method proposed in this paper solves the problem that a large amount of PD data cannot be obtained from the actual equipment, and the sample expansion algorithm makes the PD localization beyond the limits of the specific PD simulation location. Moreover, the simulation model can be changed according to different models of GIL, ending up with wider applicability. In this paper, the feasibility of the method is verified by the GIL experiment model.

II. OPTICAL SIMULATION MODEL OF GIL

A. Simulation Settings for PD Sources

In this paper, a spherical point light source placed directly under the tip of the needle-plate defect is considered a PD source. Due to the wavelength of PD light is mainly concentrated around the wavelength of green light, the simulated PD source is set to emit green light uniformly [9]. The total number of rays emitted by the simulated PD source is 250000, and the total optical radiation flux is 100W. The gas insulation medium in the GIL simulation model is SF_6 whose absorption spectrum has a negligible effect on the propagation of the PD signal [10].

In the simulation, the simulated optical signal strength and the actual PD signal strength cannot correspond to the absolute value, so this paper introduces the concept of relative light irradiance E_r to represent the light signal distribution law received by each sensor.

$$E_r = dP_r / dS \qquad (1)$$

where P_r represents the light radiation flux received by the sensor, and S represents the receiving area of the sensor.

B. Simulation Settings for GIL Material

The GIL material exerts vital influence on the propagation of PD optical signals. In this paper, the diffuse reflection model of the GIL simulation model uses bidirectional reflectance distribution function. The model represents the irradiance distribution of the reflected light in the three-dimensional hemisphere with respect to the incident light from different angles on the surface of the GIL tank with the incident point of light as the center [9].

Meanwhile, the propagation of optical signals in GIL needs to meet the following conditions:

$$\alpha + R + T = 1 \qquad (2)$$

where α represents absorption coefficient; R represents reflection coefficient; T represents diffuse reflection coefficient. For the polished and oxidized medium smooth aluminum GIL experimental tank used in the text, α=0.3, R=0.2, T=0.5.

C. Simulation Settings for the Model Size

The length of the GIL actually put into operation is long, and a large number of PD failure experiments cannot be carried out in it. Hence, this paper builds a GIL experimental tank in the laboratory to verify the feasibility of the method, as shown in Fig. 1. According to the size of the real GIL experiment tank, the simulation GIL model is set in the Tracepro software, as shown in Fig. 2. The specific size of the GIL model is listed in Table 1. Nine fully transmissive probe points are installed on the GIL, which are divided into three columns at 120° around the GIL tank.

In the process of the optical PD simulation, to change the height, the radial distance from the axis and the angle of rotation of the needle-plate defect model is to gain the optical PD information at various locations in the GIL.

Fig. 1. GIL experimental tank.
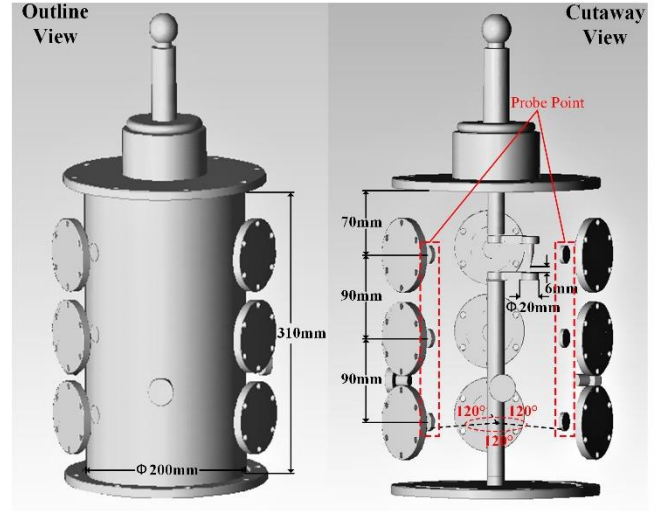
Fig. 2. GIL tank simulation structure diagram.

TABLE I. MODEL SIZE OF GIL TANK

Part	Internal height	Inner radius	Wall thickness	Inner conductor radius
Size (mm)	310	90	10	25
Part	Needle plate spacing	Needle tip length	Plate radius	Probe point radius
Size (mm)	6	25	10	10

D. Construction of Optical Simulation Fingerprint Database

Based on the simulation model mentioned above, optical PD simulation experiments are performed at N locations in the GIL simulation model, recording L_j ($j=1,2,\cdots,N$) as the location of the PD source. Each PD simulation experiment uses M detection points to detect light irradiance, recording S_i ($i=1,2,\cdots,M$) as the detection point number. When the simulated PD source is located at L_j, the irradiance of PD detected by the probe point S_i is expressed as $\varphi'_{i,j}$.

In order to avoid the impact of PD amplitude changes, this paper normalizes the PD fingerprint that is the irradiance of the light signal received by each sensor in simulation and actual measurement into [-1,1]. The rules for constructing PD fingerprints in this way are as follows:

$$\varphi_{i,j} = \frac{2 \times \left(\varphi'_{i,j} - \min\left(\varphi'_{1,j},\varphi'_{2,j},\mathrm{L},\varphi'_{M,j}\right)\right)}{\left(\max\left(\varphi'_{1,j},\varphi'_{2,j},\mathrm{L},\varphi'_{M,j}\right) - \min\left(\varphi'_{1,j},\varphi'_{2,j},\mathrm{L},\varphi'_{M,j}\right)\right)} - 1 \quad (3)$$

where $\varphi_{i,j}$ is the normalized PD fingerprint.

The optical simulation fingerprint database based on $\varphi_{i,j}$ is as follows:

$$\Psi = \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} & \mathrm{L} & \varphi_{1,N} \\ \varphi_{2,1} & \varphi_{2,2} & \mathrm{L} & \varphi_{2,N} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ \varphi_{M,1} & \varphi_{M,2} & \mathrm{L} & \varphi_{M,N} \end{bmatrix} \quad (4)$$

where M is the number of probe points, N is the number of simulation PD sources. Column vector $\Psi_j = [\varphi_{1,j},\varphi_{2,j},\cdots,\varphi_{M,j}]^T$ represents the optical simulation fingerprint of PD source L_j, which is the PD irradiance received by each simulated detection point when the PD occurs at the L_j location.

III. MACHINE LEARNING MATCHING METHOD: PSO-SVM

A. The Particle Swarm Optimization

PSO is a kind of optimization method introduced from the swarming nature of animals groups to solve the problem of optimal solution [11]. In PSO algorithm, a group of random particles is initialized to find the optimal solution through iteration. In each iteration, the particle updates itself by tracking two extreme values. One is the optimal solution of the particle itself, called individual extremum. The other is to obtain the optimal solution for the entire group, called the global extremum. When the iteration condition meets the maximum number of iterations or the optimized position reaches the threshold, the iteration stops.

Suppose there are m particles in the N-dimensional target search space to form a group. $x_i=(x_{i1},x_{i2},x_{i3},\cdots,x_{iN})$ represents the position of the ith particle in the N-dimensional search space. The flight speed is represented by vector $v_i=(v_{i1},v_{i2},v_{i3},\cdots,v_{iN})$, $i=1,2,\cdots,m$. The particle velocity and position are defined as follows:

$$\mathbf{v}_{in}^{k+1} = \omega \mathbf{v}_{in}^{k} + c_1 r_1 (\mathbf{p}_{in}^{k} - \mathbf{x}_{in}^{k}) + c_2 r_2 (\mathbf{p}_{gn}^{k} - \mathbf{x}_{in}^{k}) \quad (5)$$

$$\mathbf{x}_{in}^{k+1} = \mathbf{x}_{in}^{k} + \mathbf{v}_{in}^{k+1} \quad (6)$$

where $i=1,2,\cdots,m$ and $n=1,2,\cdots,N$. k is the number of iterations. c_1 and c_2 are the learning factor. r_1 and r_2 are the random numbers between [0,1]. ω is the inertia weight. $v_{in} \in [-v_{max}, v_{max}]$, v_{max} is set according to the situation. p_{in} is the optimal position found by the ith particle. p_{gn} is the optimal position found by the entire particle swarm.

B. The Support Vector Machine

SVM is a supervised machine learning method based on statistical Vapnik-Chervonenkis Dimension theory law [12]. It seeks to minimize the structural risk to construct the optimal hyperplane and improve the generalization ability of the learning machine. SVM minimizes empirical risk and confidence, so as to achieve a good classification effect in the case of small samples. The specific theory is as follows.

The definition of training sample set:

$$A = \{(x_i, y_i), i = 1,2,\mathrm{L},n\} \quad (7)$$

where $x_i \in R^d$, y_i is the label of x_i.

The hyperplane equation is $\omega\varphi(x)+b=0$, where ω is weight, b is threshold and $\varphi(x)$ is the nonlinear function.

The function needs to solve:

$$\min[\frac{1}{2}(w^T w)] + C\sum_{i=1}^{n} \zeta_i$$
$$s.t. \begin{cases} y_i(w^T w + b) \geq 1 - \zeta_i \\ \zeta_i \geq 0 \end{cases} \quad (8)$$

where C represents the penalty coefficient, ζ_i is the slack variable.

In order to transform the linear inseparable problem into a linear separable problem in a high-dimensional space, SVM introduces Lagrange multiplier method and KKT condition, avoiding the dimensional disaster. Then the maximization problem becomes:

$$Q(\lambda) = \sum_{i=1}^{n} \lambda_i - \frac{1}{2}\sum_{i,j=1}^{l} \lambda_i \lambda_j y_i y_j K(x_i, y_j) \quad (9)$$

The classification function is

$$f(x) = \mathrm{sgn}(\sum_{i=1}^{n} \lambda_i y_i K(x_i, x) + b) \quad (10)$$

where λ is Lagrange multiplier, $K(x_i,x)$ is inner product kernel function.

The commonly used kernel function is Radial Basis Function (RBF) kernel function:

$$K(x_i, x_j) = \exp(\frac{-\|x_i - x_j\|^2}{2\sigma^2}) \quad (11)$$

where σ is kernel function parameters.

C. PSO-SVM

Due to the fingerprint matching effect of the SVM machine learning model is closely related to the setting of its penalty parameter C and kernel function parameter σ, this paper uses the PSO algorithm to optimize these two parameters to avoid the blindness caused by artificial settings. The process of PSO optimization SVM model is as follows:
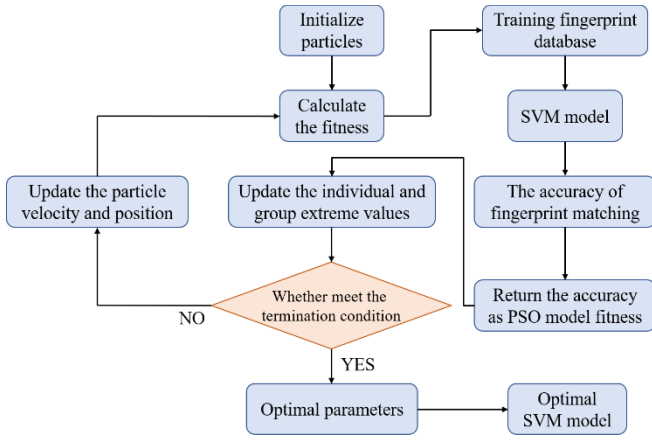
Fig. 3. PSO-SVM optimization flow chart.

IV. PD LOCALIZATION MODEL BASED ON OPTICAL SIMULATION AND PSO-SVM

This paper combines the optical PD simulation and the PSO-SVM machine learning algorithm to solve the problem that the actual PD test cannot be carried out on site. Moreover, through the simulation model, a localized fingerprint database with richer samples can be constructed, thereby improving the accuracy of PD localization. The whole localization mainly includes two parts. One is the construction of the optical PD simulation fingerprint database. The other is the PD fingerprint matching. The overall flow of the PD localization method proposed in this paper is shown in Figure 4.
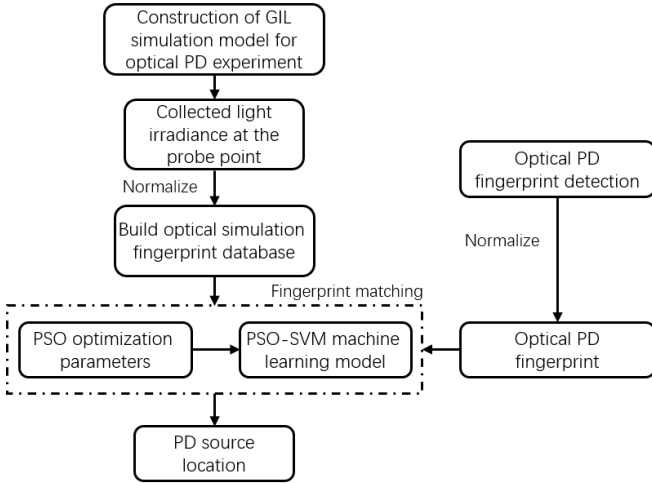
Fig. 4. PD localization flowchart.

A. Construction of Optical PD Simulation Fingerprint Database

We select PD sources as many as possible and distribute them as evenly as possible in each area of the simulation tank. To further expand the fingerprint scale of the fingerprint database, the biharmonic spline interpolation method [13] is used to interpolate the PD simulation sources. Then we select a certain number of fingerprints in light of the situation to establish the final PD fingerprint database.

B. PD Fingerprint Matching

The PSO-SVM model is trained with the constructed optical PD simulation fingerprint database as a training set. Then, the detected optical PD fingerprints are put into the optimized and trained PSO-SVM model for matching. The coordinate of the PD source corresponding to the matched PD fingerprint is the actual location of the PD source.

V. EXPERIMENT VERIFICATION

A. Construction of Optical PD Simulation Fingerprint Database

Based on the above GIL simulation model, a cross section is selected every 10 mm from top to bottom inside the tank body, and a total of 27 cross sections are selected. Each cross section is divided by the radius of every 30°, for a total of 12 radii. Then, at each radius, points at distances of 0mm, 24mm, 44mm, 64mm, and 84mm from the center of the circle are selected for optical PD simulation experiments. As a result, a total of 27×12×5=1620 PD simulation experiments are conducted in the simulation tank, where the simulation position of the PD source on each cross-section is shown in Figure 5.
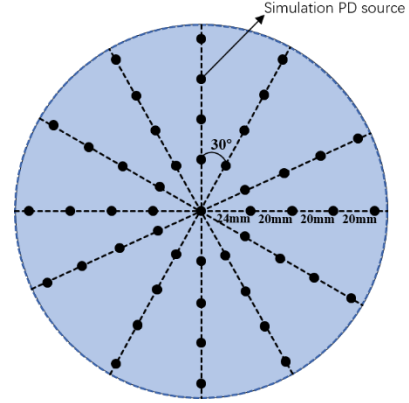
Fig. 5. Simulation location of the PD source on each cross section.

The biharmonic spline interpolation method is performed on the light irradiance of 1620 simulated PD sources. Given that machine learning model training cannot directly input all the interpolation data, considering the matching efficiency, we select 8370 PD fingerprints to form an optical simulation fingerprint database. Taking a simulated probe point near the top of the tank as an example, the light irradiance distribution of each PD source on the 27 cross-sections collected is shown in Figure 6, which is the PD simulation fingerprint map of this probe point.
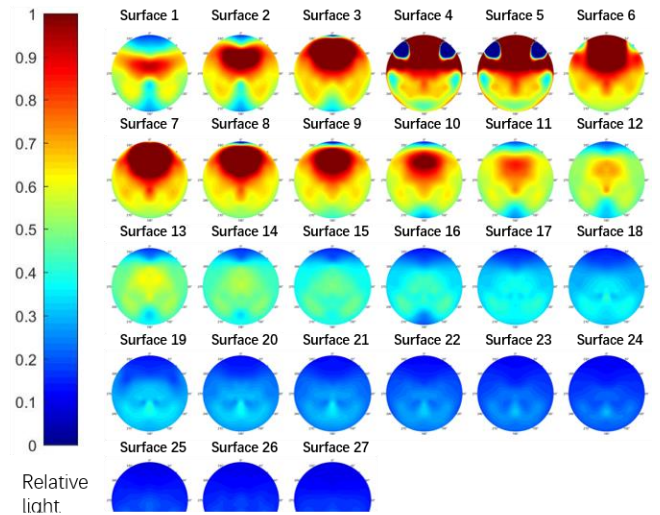
Fig. 6. Fingerprint maps on 27 cross-sections by the probe point near the top.

B. PD Fingerprint Matching

In order to verify the feasibility of the method proposed in this paper, we built a GIL PD localization detection platform in the laboratory, as shown in the Figure 7.
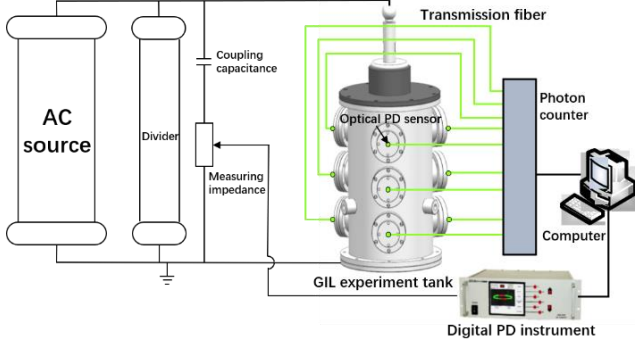
Fig. 7.   Experiment platform for PD localization detection.

The experiment platform mainly consists of corona-free AC voltage regulation system, GIL experimental tank, optical sensor based on fluorescent fiber, digital PD instrument (Hafley DDX 9121b) and photon counter (HAMAMATSU H11890-210).

In the experiment, the PD at different locations is achieved by adjusting the angle, height and radial length of the needle plate model in the GIL experiment tank. In this paper, 12 typical PD sources are selected for experiment. Before each experiment, the tank was sealed and evacuated, and then filled with 0.2MPa SF_6 gas. Finally, the voltage is increased slowly until a clear and stable PD signal appears, recording the number of photons received by the 9 optical sensors.

When using the photon counter to collect, in order to reduce the impact of fluctuations in the PD signal strength, we collect 60 threshold photons for each PD. Then take the average value of 60 thresholds as the light irradiance value detected by the sensor. The light irradiance values of the nine optical sensors are normalized to construct the actual optical PD fingerprint of one PD source.

Before fingerprint matching, the PSO algorithm is used to optimize the penalty parameter C and the kernel function parameter σ of the SVM. The parameters of the PSO algorithm are set as: the group number is 20, learning factor $c_1=c_2=2$, the maximum number of iterations is 150. Through the optimization of PSO, the penalty parameter $C=0.12$ and the kernel function parameter $\sigma=0.54$. Finally, the optimized SVM model is used to match the actual optical PD fingerprints.

C. Result and Analysis

According to the above experimental process, the localization results of 12 PD test sources are shown in Figure 8. Taking the size of the experimental tank into account, these 12 experimental points are randomly selected from typical space areas in the tank, which can fully verify the effectiveness of this method.

In Figure 8, the 12 PD sources are numbered 1 to 12 according to the value of the localization error. The experiment results show that the average error of PD localization in the experimental GIL using the method proposed in this paper is 10.58mm. The minimum localization error and the maximum localization error are 1.00mm and 21.36mm, respectively. The error of 91.67% is less than 20mm, which basically meets the localization requirements in GIL. Moreover, in the matching process of PD fingerprints, the identification of fingerprint features is based on the difference

in the intensity of light irradiation received by each sensor. In this paper, although the size of the tank in the experiment is smaller than the actual GIL, the difference in light irradiance between the sensors in the small tank is relatively minute. The difficulty of identification and matching is greater than that of the actual GIL. Therefore, the above verification result can be achieved in the experimental small GIL in this paper, and the fingerprint matching result in the actual large-scale GIL will be better.

Compared with the existing PD localization method, the average localization result of PD using the UHF method in [14] and [15] is 125mm and 89mm respectively. The average localization result of PD using the ultrasonic localization is 17.29mm in [16]. The average localization result of PD using the Fiber-Optic acoustic sensor array is around 50mm in [17]. In summary, the localization accuracy of the method in this paper is higher than that of the current typical PD localization methods, which can reach an error of about 10mm.
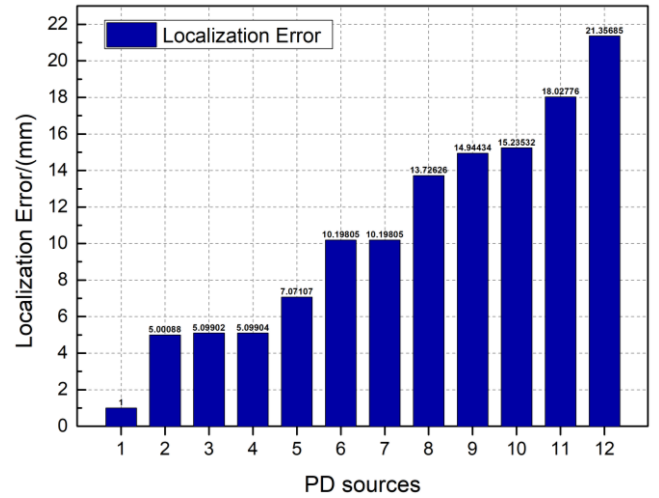
Fig. 8.   Experimental result of PD localization.

VI. CONCLUSION

The GIL PD localization method proposed in this paper introduces optical simulation data into the field of PD localization. Combined with PSO-SVM machine learning algorithm, the optical fingerprint is matched to obtain the location of PD sources. Through relevant experiments, the feasibility of the method is verified, and the following conclusions are drawn:

(1) By establishing GIL simulation model in Tracepro software and conducting optical PD simulation experiment in it, the difficulty of establishing PD fingerprint database through field experiments is avoided, which improves detection efficiency.

(2) This paper interpolates the optical simulation data to expand the sample density of the fingerprint database. As a result, the range that the fingerprint database can locate is expanded, and the accuracy of fingerprint localization is improved.

(3) In this paper, the combination of optical simulation fingerprint database and PSO-SVM machine learning algorithm achieves an average localization error of 10.58mm in the experimental GIL tank. It shows that this method can effectively guide the maintenance of PD faults in GIL.

REFERENCE

[1] M. Tenzer, H. Koch, and D. Imamovic, "Underground transmission lines for high power AC and DC transmission," in 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), 2016, pp. 1-4.

[2] H. Okubo, M. Yoshida, T. Takahashi, T. Hoshino, M. Hikita, and A. Miyazaki, "Partial discharge measurement in a long distance SF/sub 6/ gas insulated transmission line (GIL)," IEEE Transactions on Power Delivery, vol. 13, no. 3, pp. 683-690, 1998.

[3] A. S. Kumar, R. P. Gupta, K. Udayakumar, and A. Venkatasami, "Online partial discharge detection and location techniques for condition monitoring of power transformers: A review," in 2008 International Conference on Condition Monitoring and Diagnosis, 2008, pp. 927-931.

[4] Y. Liu et al., "Novel and necessary assessment procedure ensuring UHF partial discharge on-line supervision system effective for GIS/GIL," in 2017 4th International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST), 2017, pp. 803-807.

[5] H. Hou, G. Sheng, and X. Jiang, "Localization Algorithm for the PD Source in Substation Based on L-Shaped Antenna Array Signal Processing," IEEE Transactions on Power Delivery, vol. 30, no. 1, pp. 472-479, 2015.

[6] S. M. Markalous, S. Tenbohlen, and K. Feser, "Detection and location of partial discharges in power transformers using acoustic and electromagnetic signals," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 15, no. 6, pp. 1576-1583, 2008.

[7] S. Biswas, C. Koley, B. Chatterjee, and S. Chakravorti, "A methodology for identification and localization of partial discharge sources using optical sensors," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 19, no. 1, pp. 18-28, 2012.

[8] H. Guzman and H. Kranz, "Analysis of internal partial discharges in solid dielectrics using an electro-optical measurement system," in Proceedings of 1995 IEEE 5th International Conference on Conduction and Breakdown in Solid Dielectrics, 1995, pp. 239-243.

[9] S. Biswas, D. Dey, B. Chatterjee, and S. Chakravorti, "An approach based on rough set theory for identification of single and multiple partial discharge source," International Journal of Electrical Power & Energy Systems, vol. 46, pp. 163-174, 2013.

[10] M. Ren, B. Song, T. Zhuang, and S. Yang, "Optical partial discharge diagnostic in SF6 gas insulated system via multi-spectral detection," ISA transactions, vol. 75, pp. 247-257, 2018.

[11] Y. Xu, Y. Qian, G. Sheng, X. Jiang, X. Zhou, and Z. Wang, "Simulation analysis on the propagation of the optical partial discharge signal in I-shaped and L-shaped GILs," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 25, no. 4, pp. 1421-1428, 2018.

[12] W. Zihao, L. Lan, X. Zongyi, and C. Guangtao, "The forecasting model of wheelset size based on PSO-SVM," in 2019 Chinese Control And Decision Conference (CCDC), 2019, pp. 2609-2613.

[13] Z. Qian, C. Zhou, J. Cheng, and Q. Wang, "Identification of conductive leakage signal in power cable based on multi-classification PSO-SVM," in 2017 IEEE 5th International Symposium on Electromagnetic Compatibility (EMC-Beijing), 2017, pp. 1-4.

[14] D. T. Sandwell, "Biharmonic spline interpolation of GEOS‐3 and SEASAT altimeter data," Geophysical research letters, vol. 14, no. 2, pp. 139-142, 1987.

[15] Z. Li, L. Luo, Y. Liu, G. Sheng, and X. Jiang, "UHF partial discharge localization algorithm based on compressed sensing," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 25, no. 1, pp. 21-29, 2018.

[16] T. Lijun, L. Richeng, D. Min, and S. Jun, "Study of Partial Discharge Localization Using Ultrasonics in Power Transformer Based on Particle Swarm Optimization," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 15, no. 2, pp. 492-495, 2008.

[17] C. Gao, L. Yu, Y. Xu, W. Wang, S. Wang, and P. Wang, "Partial Discharge Localization Inside Transformer Windings via Fiber-Optic Acoustic Sensor Array," IEEE Transactions on Power Delivery, vol. 34, no. 4, pp. 1251-1260, 2019.

The 500kV Oil-filled Submarine Cable Temperature Monitoring System Based on BOTDA Distributed Optical Fiber Sensing Technology

Yu Chen*
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenyu@xjtu.edu.cn

Shuang Wang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
shuang@xjtu.edu.cn

Yi Hao
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
haoyi187009@stu.xjtu.edu.cn

Kai Yao
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
yaokai17@stu.xjtu.edu.cn

Hanzhi Li
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
lhz3119304518@stu.xjtu.edu.cn

Feng Jia
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
jia2019@stu.xjtu.edu.cn

Qingyun Shi
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
sqy3119304516@stu.xjtu.edu.cn

Dongli Yue
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
dongliyue@stu.xjtu.edu.cn

Yonghong Cheng
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
cyh@xjtu.edu.cn

Abstract—The 500 kV oil-filled AC submarine cable in Hainan Networking System is the first large capacity, ultra-high voltage cross-sea submarine power cable in China, which is 31 kilometers long and bundled with a submarine optical cable. And temperature monitoring on the submarine cable is of great significance to control the current rating and ensure the operation reliability. Based on the structural characteristics and setting condition of this submarine cable, a temperature monitoring system is established in this paper based on the Brillouin optical time-domain analysis (BOTDA) technology. In this system, the optical cable is used to monitor the temperature of the submarine cable surface, and an optimized thermal circuit model is proposed to calculate the internal temperature distribution. Moreover, an onshore simulation experimental platform is established on the same cable as the Hainan Networking System for the first time, which proves the accuracy of the temperature monitoring system and the calculation method.

Keywords—500 kV, oil-filled submarine cable, temperature distribution, onshore simulation test, thermal circuit model

I. INTRODUCTION

The high voltage submarine cables, as the main power transmission channel between islands and harbors, have been widely developed, and many countries around the world have carried out various studies on them [1]. The oil-filled 500 kV submarine cable of China's Hainan Networking System extends from Guangdong Province to Hainan Province and spans the Qiongzhou Strait, which has significant economic and social value [2]. The temperature of the submarine cable while it is in operation is an important parameter to determine its loading and evaluate its operational status [3]. As electricity demand continues to increase, the load also increases, which also leads to accretion in the operation temperature of the cable. However, excessive load and temperature will accelerate the aging of the insulating material and even cause

*Yu Chen is the corresponding author. (e-mail: chenyu@xjtu.edu.cn).

thermal breakdown, affecting the safe and stable operation of the power system. At the same time, the insulation faults of high voltage cables are usually manifested by local temperature changes or abnormalities [4]. Therefore, establishing a temperature online monitoring system of the submarine cable is significant to clarify its loading and ensure its safe operation.

The 500 kV oil-filled submarine cable in the Hainan Networking System can be divided into four parts according to different installation environments: the air part, the landing part, the intertidal part, and the seabed part [5]. Due to the complex and changeable environment of this submarine cable, it is impossible to test its temperature distribution directly. With the unique structure of the submarine cable, that bundled with optical cable, the distributed optical fiber sensing technology based on BOTDA can be used to deduce the temperature change of the optical unit by monitoring the Brillouin frequency shift in optical cable [6], and thus obtain the temperature of the submarine cable surface. And based on this, the temperature distribution of each layer inside the submarine cable can be calculated with the thermal circuit model of cables. The commonly used thermal circuit model IEC60287 is proved an effective method to derive the temperature of the cable conductor from environment temperature [7]. It simplifies the actual cable structure into a one-dimensional thermal circuit model and derives the temperature of the main structure of the cable layer by layer. However, since the optical cable in this testing system directly tests the temperature of the submarine cable surface, which is different from the environment temperature in the IEC60287 model, thus the thermal circuit model needs to be optimized according to the test conditions in this paper.

In this paper, a temperature monitoring system is established for the 500 kV oil-filled submarine cable in Hainan Networking System, which is relying on the bundled optical cable and the BOTDA technology. And the IEC60287 thermal circuit model is optimized to calculate the temperature

distribution inside the submarine cable based on the surface temperature achieved with optical cable. Moreover, an onshore simulation experimental platform is established that uses the same cable as the Hainan Networking System for the first time. By testing the temperature of each layer in the 500 kV oil-filled submarine cable with different loads, the optimized thermal circuit model in this paper is verified.

II. ONLINE TEMPERATURE MONITORING SYSTEM OF THE 500 kV SUBMARINE CABLE

The submarine cable and optical cable of the Hainan Networking System have been in operation for several years, and the online temperature monitoring system needs to be designed according to their setting condition. The entire length of the 500 kV oil-filled submarine cable is bundled with submarine optical cable, with the bundle way shown in Fig. 1. The type of the optical cable is URC-I-G24QEVA-φ3.7-R1.9/2.4, of which the cross-sectional structure is also shown in Fig. 1. Every optical cable has 24 single-mode G.652D optical fibers in the middle. The structure diagram of the submarine cable is shown in Fig. 2, with the scales and materials of each layer summarized in TABLE I.

(a)   (b)
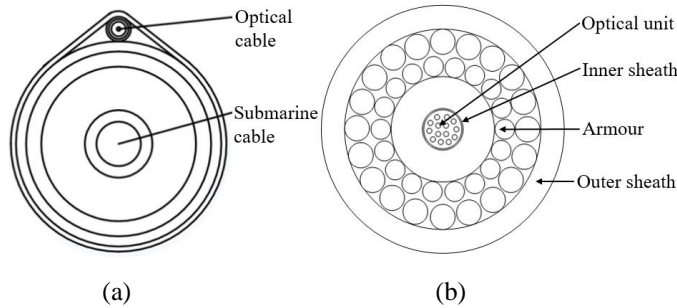
Fig. 1.   (a) the structure diagram of 500 kV submarine cable bundled with a submarine optical cable and (b) the structure diagram of the optical cable.
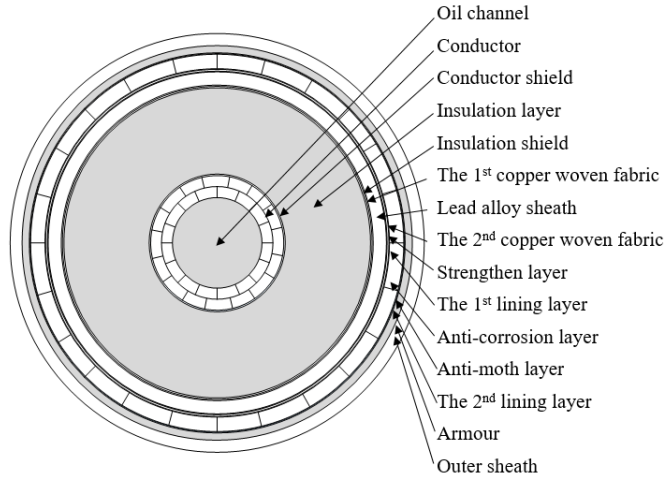
Fig. 2.   The structure diagram of the submarine cable.

TABLE I.   THE STRUCTURES AND MATERIALS OF EACH LAYER OF THE 500 kV SUBMARINE CABLE

Layers	Diameter （mm）	Thickness (mm)	Material
Oil channel	30	15	Cable oil
Conductor	44.6	7.3	Copper
Conductor shield	45.6	0.5	Carbon black paper
Insulation layer	102.7	28.55	Impregnated paper tape
Insulation shield	103.7	0.5	Carbon black paper
Copper woven fabric	104.5	0.4	Copper
Strengthen layer	115	0.6	Bronze
The 1st lining layer	115.4	4	Polypropylene
Anti-corrosion layer	125	5	Polyethylene
Anti-month layer	125.4	0.2	Copper
The 2nd lining layer	125.9	4	Polypropylene
Armour	130.7	2.4	Copper
Outer sheath	138.7	4	Polypropylene

According to the characteristics of the submarine cable, the optical cable, and the setting method of these two, the temperature online monitoring system set up for the Hainan Networking System in this paper is as shown in Fig. 3. Two optical fibers in the optical cable are selected as temperature sensors. The ends in the Hainan terminal station are connected to the DITEST STA-R Series Brillouin Optical Time Domain Reflectometry (DTS) to test the Brillouin frequency shift by temperature in the optical fibers. And the other ends in Guangdong terminal station are connected to form a loop. While testing, one pulse light, and one testing light are input into the two optical fibers, separately. When the frequency difference of these two lights is equal to the Brillouin frequency shift of a certain area in the optical fiber, Brillouin scattering will occur, and the energy will be transferred between these two lights. By detecting the power of the output testing light, the frequency difference corresponding to the maximum transfer energy in each area of the optical fiber can be determined, thereby realizing the distributed temperature monitoring.

III. THE EQUIVALENT THERMAL CIRCUIT MODEL FOR THE 500 kV OIL-FILLED SUBMARINE CABLES

With the above submarine cable temperature monitoring system, the surface temperature of the submarine cable can be achieved from time to time. On this basis, the temperature distribution of the submarine cable can be derived from the surface temperature by the optimized thermal circuit model IEC60287. The classic IEC60287 thermal circuit model connects the thermal resistance of each layer of the cable in series, thereby equating the heat transfer path of the cable to a one-dimensional thermal circuit model to calculate the temperature difference between the environment and the conductor.
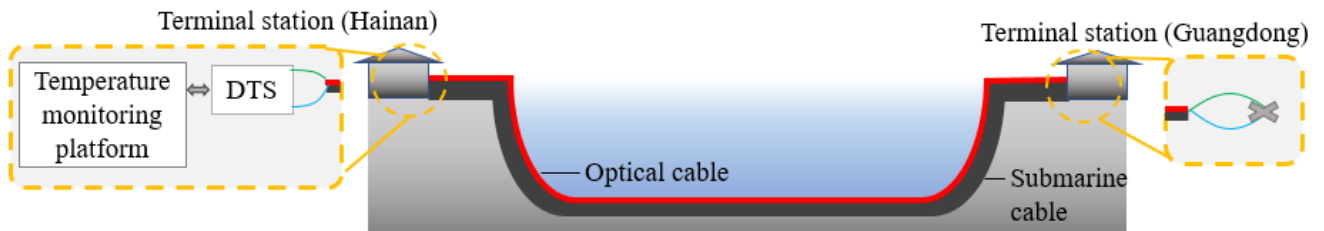
Fig. 3.   The schematic diagram of the temperature monitoring system of the submarine cable in the Hainan Networking System.

The thermal circuit model is as shown in Fig. 4. And the conductor temperature can be calculated by Eq. (1).

$$\theta_c = \theta_a + (I^2 + 0.5W_d)T_1$$
$$+[I^2R(1+\lambda_1)+W_d]T_2 \qquad (1)$$
$$+[I^2R(1+\lambda_1+\lambda_2)+W_d](T_3+T_4)$$

Where θ_c is the conductor temperature (ºC), θ_a is the environment temperature (ºC), W_d is the dielectric loss (W/m), R is the resistance per kilometer (Ω/km), T_1 is the thermal resistance of the conductor (K•m/W), T_2 is the thermal resistance of the armor (K•m/W), T_3 is the thermal resistance of the outer sheath (K•m/W), T_4 is the thermal resistance of the environment (K•m/W), λ_1 is the loss factor of the armor, λ_2 is the loss factor of the outer sheath.
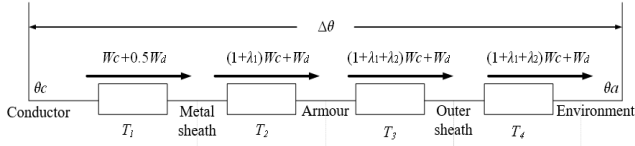
Fig. 4.   The IEC60287 thermal circuit model.

To make the calculation more accurate, the above mentioned IEC60287 thermal circuit model needs to be optimized according to the test system. The test object of the temperature monitoring system in this paper is the surface temperature of the cable, and the key structures inside the submarine cable are mainly the outer sheath, the armor, the lead alloy sheath, and the conductor. Therefore, the thermal circuit between the conductor and the surface of the submarine cable can be equivalent to a series of three thermal resistances, the lead alloy sheath and armor are the thermal circuit boundary layers, and between the boundary layers are the equivalent thermal resistance. The thermal circuit model of the submarine cable is as shown in Fig. 5. So the calculation equation for the conductor temperature can be optimized as Eq. (2).

$$\theta_c = \theta_o + \left(I^2R + 0.5W_d\right)T_1$$
$$+\left[I^2R(1+\lambda_1)+W_d\right]T_2 \qquad (2)$$
$$+\left[I^2R(1+\lambda_1++\lambda_2)+W_d\right]T_3$$

Where θ_c is the conductor temperature (ºC), θ_o is the temperature of the submarine cable surface (ºC), W_d is the dielectric loss (W/m), T_1 is the thermal resistance between the conductor and the lead alloy sheath (K•m/W), T_2 is the thermal resistance between the lead alloy sheath and the armor (K•m/W), T_3 is the thermal resistance between the armor and the outer sheath (K•m/W), λ_1 is the loss factor of the lead alloy sheath, λ_2 is the loss factor of the armor. The values of these parameters can be calculated according to the scales and materials of each layer listed in TABLE I.
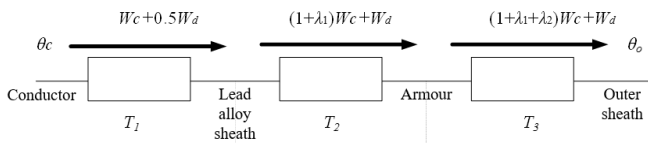
Fig. 5.   The thermal circuit model of the submarine cable.

IV. ONSHORE SIMULATION EXPERIMENTS

Due to the complex and changeable operation environment of the submarine cable, it is impossible to measure its surface temperature directly, and it is also impossible to conduct destructive experiments on the submarine cables in operation to test its internal temperature. Thus, to verify the accuracy of the above calculation model, this paper sets up onshore simulation experiments in the laboratory and uses the same submarine cable and optical cable as the Hainan Networking System for the first time.

A. Experimental Materials and Equipment

The 500 kV oil-filled submarine cable is provided by the China Southern Power Grid, with a total length of about 1200 mm. To test the temperature distribution of the key layers inside the submarine cable, this paper drills eight holes with different depths at different positions in the middle of the submarine cable, and penetrated the middle of the conductor, the insulation layer, the lead alloy sheath, and the outer sheath, respectively. The diameters of the holes are all 2 mm. The positions of the holes are as shown in Fig. 6. The current generator is provided by the Shandong Taikai High Voltage Switchgear Co., LTD, which can generate large current up to 1600 A. The temperature monitoring device is Tektronix DAQ6510, with the accuracy for temperature testing is 0.01ºC.
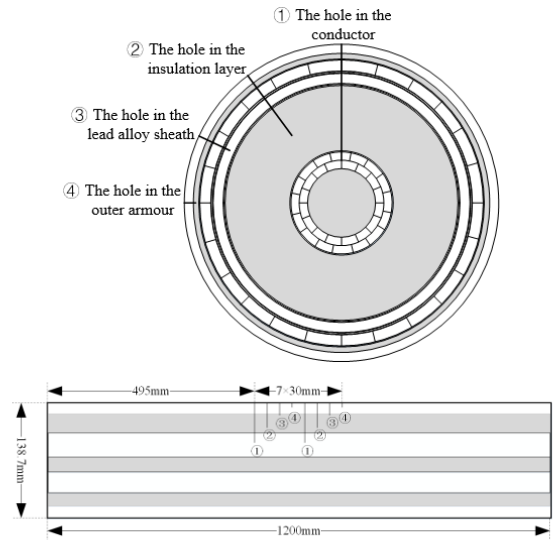
Fig. 6.   The schematic diagram of the drilling position on the 500 kV oil-filled submarine cable.

B. Experimental Setup

The experimental setup is as shown in Fig. 7. The current generator is connected to both ends of the submarine cable to simulate the load when the submarine cable is in operation. There are six thermocouples during the test, four of which are inserted into the corresponding holes of different key layers, and the other two are attached to the surface of the submarine cable and suspended in the air, separately. All the thermocouples are connected to the temperature monitoring device and further connected to the PC to monitor and record the temperature.
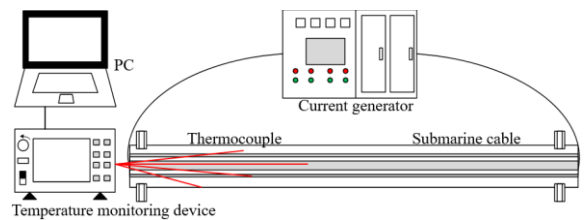
Fig. 7.   The schematic diagram of the temperature rise experiment of the 500 kV oil-filled submarine cable.

While conducting the temperature rise experiments, set the load currents of 300A, 600A, 900A, and 1200A, respectively, and observe the temperature changes of the key layers of the submarine cable. When the temperature is stable, record the temperature of each layer of the submarine cable.

V. RESULTS AND DISCUSSION

A. The Results Achieved From the Onshore Experiments

The relationship between the temperature of different layers in the 500 kV oil-filled submarine cable and the load current achieved from the temperature rise experiment is shown in Fig. 8, with the specific values listed in TABLE II. The higher the load, the higher the temperature of each layer. And the temperature of the structure closer to the center is more sensitive to the load changes.
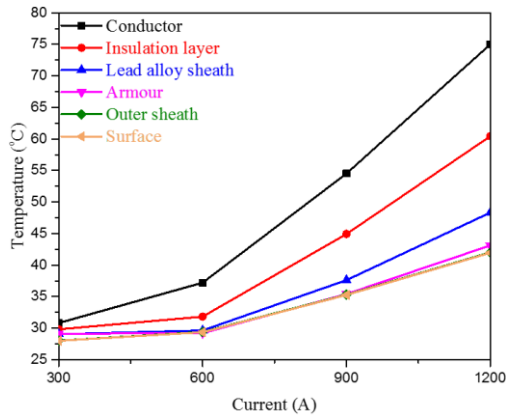
Fig. 8.   The experimental results of the temperature of the key layers in the 500 kV oil-filled submarine cable at different current loads.

TABLE II.   THE TEMPERATURE OF THE KEY LAYERS IN THE SUBMARINE CABLE ACHIEVED FRON EXPERIMENTS

Layers	Temperatures at different load (ºC)			
	300 A	600 A	900 A	1200 A
Conductor	30.82	37.22	54.52	75.01
Insulation layer	29.83	31.83	44.93	60.43
Lead alloy sheath	29.04	29.64	37.62	48.33
Armour	29.13	29.22	35.41	43.12
Outer sheath	28.02	29.41	35.32	42.01
Surface	28	29.36	35.27	41.95

B. The results Achieved From the Thermal Circuit Model

According to the structural dimensions and the corresponding material characteristics of each layer in the 500 kV oil-filled submarine cable given in TABLE I, the parameter values of the thermal circuit model are calculated and listed in TABLE III.

TABLE III.   THE THERMAL CIRCUIT MODEL PARAMETERS OF THE 500 kV SUBMARINE CABLE

Parameters	Values
R (Ω/km)	0.0286
T_1 (K·m/W)	0.6787
T_2 (K·m/W)	0.0559
T_3 (K·m/W)	0.0567
λ_1	0.1
λ_2	0.491

Introduce the parameter values and the surface temperature values in TABLE II into Eq. (2). The calculated conductor temperature at different loads is summarized in

TABLE IV. By comparing the calculation results with the conductor temperature obtained from the temperature rise experiments, it can be seen that the optimized thermal circuit model is very accurate, and the error between the two is less than 2.26%.

TABLE IV.   THE CONDUCTOR TEMPERATURE ACHIEVED FROM THE OPTIMIZED THERMAL CIRCUIT MODEL

Current (A)	Conductor temperature (ºC)		Error（%）
	Calculation result	Experimental results	
300	30.14	30.82	2.26
600	37.91	37.21	1.82
900	54.51	54.52	0.02
1200	76.15	75.01	1.5

VI. CONCLUSION

According to the structural characteristics and setting condition of the 500 kV oil-filled submarine cable bundled with an optical cable in the Hainan Networking System, a temperature monitoring system based on BOTDA technology is designed and established at the operating site. In this system, the optical cable monitors the temperature of the submarine cable surface. And the IEC 60287 thermal circuit model is optimized based on the temperature monitoring system to calculate the internal temperature distribution. Besides, an onshore simulation experimental platform is established in the laboratory, which conducted the simulation tests on the same submarine cable like the ones in the Hainan Networking System for the first time. By comparing the conductor temperature at different load achieved from experiments and the calculation, the accuracy of the optimized thermal circuit model is proved, and the error between the two is less than 2.26%. Based on this result, the above optimized thermal circuit model is applied to the temperature monitoring system of the 500 kV oil-filled submarine cable, and the actual operation and maintenances data will be analyzed in the future researches.

REFERENCES

[1]   Y. Wang, "Review on submarine cable projects for power transmission worldwide," Southern Power System Technology, vol. 6, pp. 26-30, 2012.

[2]   X. Huang, Y. Zang, J. Wang, "Study on 500kV submarine cable temperature online monitoring of Hainan Grid Interconnection System," Electric Wire & Cable, pp. 24-27, 2014.

[3]   W. Zhou, X. Wang, J. Bai, "Comparative Analysis on the carrying Capacity and Temperature Rise of New and Old 35kV Conductors," Southern Power System Technology, vol. 8. pp. 92-95, 2014.

[4]   Y. Tian, J. Guo, M. Fu, S. Hou, B. Hui, R. Zhuo, "Review on power transmission projects with UHVAC XLPE cables," Southern Power System Technology, vol. 8, pp. 30-36, 2016.

[5]   L. Wei, C. Yan, X. Huang, Z. Cen, W. Zhang, "Analysis and discussion on change regulation about the operating temperature of 500 kV submarine cable," Electric Power Information and Communication Technology, vol. 15, pp. 6-11, 2017.

[6]   X. Huang, K. Yao, G. Sun, J. Zhou, Y. Chen, Q. Guo, etc., "Onshore simulation test setup for temperature and strain sensing of submarine cable based on BOTDA," 2nd IEEE International Conference on Electrical Materials and Power Equipment-ICEMPE2019. Guangzhou, China. April, 2019.

[7]   J. Zhou, K. Yao, X. Huang, G. Sun, W. Zhang, A. Ashtaq, etc., "Temperature calculation and measurement on power cable conductor based on equivalent thermal circuit and BOTDA," 19th IEEE Conference on Environment and Electrical Engineering-EEEIC2019. Genoa, Italy. June, 2019.

Stator Inter-turns Short Circuit Fault Detection in DFIG Using Empirical Mode Decomposition Method on Leakage Flux

Attiq Ur Rehman
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
atnutkani@stu.xjtu.edu.cn

Yu Chen*
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenyu@xjtu.edu.cn

Guorui Huang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
huang0923@stu.xjtu.edu.cn

Yan Yang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
yangyan0415@stu.xjtu.edu.cn

Shuang Wang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
shuang@xjtu.edu.cn

Yihan Zhao
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
zhaoyihan@stu.xjtu.edu.cn

Yong Zhao
Xi'an Thermal Power Research
Institute Co. Ltd
Xi'an, China
zhaoyong@tpri.com.cn

Yonghong Cheng
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
cyh@xjtu.edu.cn

Toshikatsu Tanaka
IPS Research Center
Waseda University
Fukuoka, Japan
t-tanaka@waseda.jp

Abstract—**Inter-turn short-circuit fault (ITSCF) in doubly fed induction generator (DFIG) effect the performance of the wind turbine, so the surveillance of DFIGs is of very importance for the safe operation of the wind turbine. To solve this issue, an experimental setup was constructed for the detection of stator ITSCF. This experimental setup consist of three-phase wound rotor induction machine (WRIM), monitoring systems, frequency converters and related control and a drive motors. This experimental system can truly simulate the actual working conditions of the machine when the machine is short-circuited, and the experimental simulation can be closer to the operation process of the double-fed induction generator. Flux monitoring method was used to detect the ITSCF in stator windings. Time domain analysis was used on flux signal to detect the ITSCF at its initial stage of the fault. Empirical Mode Decomposition (EMD) method was applied on flux to detect the stator ITSCF. The EMD of the flux signal in the stator winding ITSCF yields 9 layers of empirical modal components IMFs. The IMF5 and IMF6, was found useful to discriminate the normal and faulty condition of the machine. The results show that EMD analysis on flux monitoring is advance and effective way than the other methods to detect the ITSCF in stator windings.**

Keywords—DFIG, empirical mode decomposition method, inter-turn short circuit fault, leakage flux, stator windings

I. INTRODUCTION

Nowadays wind power generation is playing a vital role in the energy sector and is getting more and more attention. Wind turbines can use different kinds of generators, for example wound rotor induction generators, squirrel-cage induction generators and synchronous generators [1]. Squirrel-cage induction generators run at a constant speed, and changes in wind speed may the source of generator failures. Wind turbine systems based on wound rotor induction generators (recognized as doubly-fed induction generators (DFIG)) are popular because of their advantages: low cost, reduced losses efficiency, low inverter ratings, active and reactive power control and higher costs [2]. Therefore, it is widely used in wind turbine.

Wind farms are usually situated in remote place and poor-environment areas and are commonly installed 50 to 80 meters above the ground. All these features put difficulties in maintenance work and increase maintenance costs. The report of Electric Power Research Institute pointed out that 37% of machine failure is due to stator related faults [3]. From these 37% stator faults most faults are allied to stator windings faults which are caused by inter-turn short-circuit fault (ITSCF). If ITSCF didn't detect at its initial stage then it grow to serious faults [4]. Therefore, it is important to monitor the machine on continues based to detect the potential failures that could reduce the losses and the maintenance cost.

Several data related to condition monitoring has been surveyed with various aims, for instance, practical approaches regarding condition monitoring [5], or other technical problems associated to them [6]. Practical Implementation of condition monitoring systems will equip the system to detect the faults at initial stages and will result in eliminating or significantly reducing the costs responsible for the unplanned and unscheduled repairing. Techniques offered through various applications are vibration monitoring, torque monitoring, temperature monitoring, acoustic monitoring, electrical signals monitoring and oil monitoring. The electrical signal monitoring is widely used in CMS, which includes electrical current, voltage, power and flux monitoring.

After selecting the monitoring method, the process of the signal monitoring is crucial. There are various options to choose from to apply the signal processing techniques to detect the faults. These can be categorized as either the frequency domain or time domain or the time-frequency [7]. In the time-domain method, an analysis of the time waveform is carried out by using statistical approaches such as measuring the mean, or the RMS (root mean square) value etc. [8]. Whereas in the frequency-domain method analysis is carried out on the transformation of the waveform signal existing in the time domain into the frequency domain. The

---
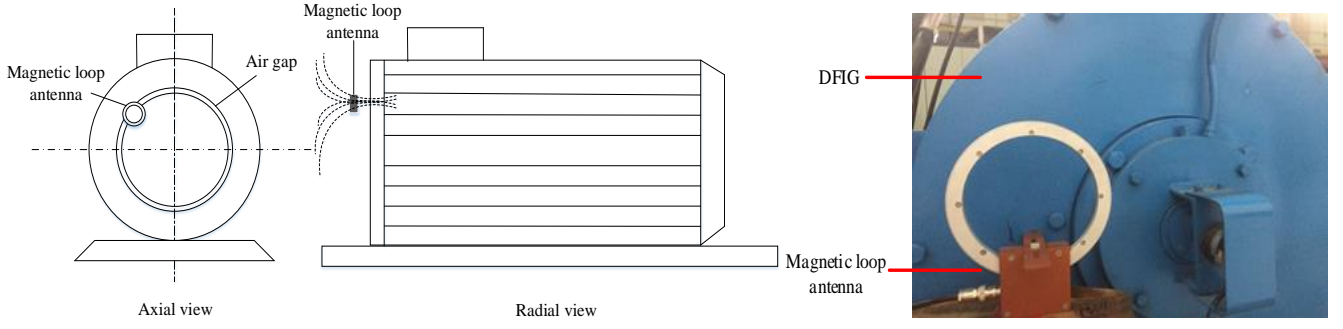
Fig. 1. (a) Schematic diagram of magnetic loop antenna test location, (b) Experimental scheme of magnetic flux leakage signal acquisition

typical method includes the spectrum analysis by using Fast Fourier Transform (FFT). On the contrary, the frequency-domain analysis has limitations such as it limit in stationary waveform signals. An empirical mode decomposition (EMD) is new time-frequency based analysis method, which was designed by Huang and his colleagues [9], which has the attractive feature of robustness in the presence of nonlinear and non-stationary data. EMD is to adaptively decompose the signal into a series of frequency components close to a single empirical mode function (IMF) according to the characteristics of the signal. Hilbert spectrum can gain the instantaneous frequency and instantaneous amplitude of the empirical mode component of each layer, analyzing the current signal and observe the change of IMF amplitude.

II. EXPERIMENTAL SETUP

The experimental setup uses a 100 kW of induction machine with a drive motor. The function of the drive motor is to rotate the induction machine. To decrease the harmonics, two back-to-back PWM VSI converters were used. To perform the ITSCFs and control the speed of the machine a control system was used. A short circuit cabinet connects the control system and the induction machine. To obtain the real time data, a condition monitoring system was used. The experiment was performed many times to verify the validity of the data.

To effectively study the magnetic flux leakage signals at different positions of the generator, magnetic loop antennas were placed at the axial positions of the generator, and the magnetic leakage of the generator at axial positions was collected through the signal, the schematic diagram of the test position of the magnetic loop antenna is shown in Fig. 1 (a). The magnetic loop antenna is placed 7cm away from the generator casing, and the magnetic loop antenna is placed parallel to the generator casing to perform the leakage magnetic flux measurement of the axial positions of the generator, the physical diagram of the test position of the experimental system is shown in Fig. 1 (b).

TABLE I.   RMS VALUES OF FLUX UNDER DIFFERENT NUMBER OF SHORTED TURNS

Shorted turns	RMS
0	0.001121
2	0.002226
3	0.002525
4	0.005711
5	0.010284
6	0.013824
7	0.014387
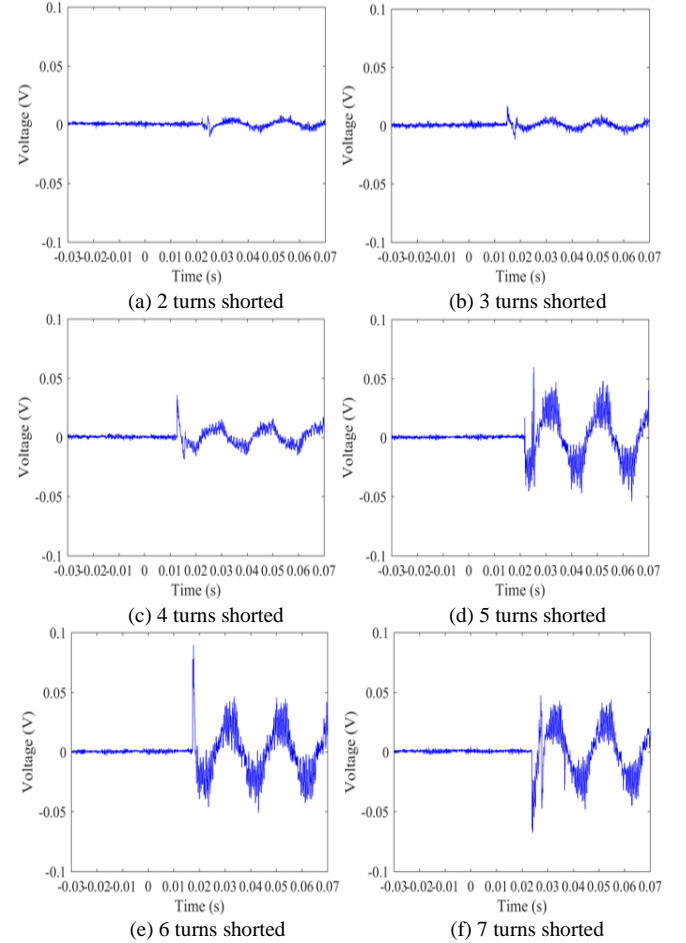
Fig. 2. Flux analysis for stator ITSCF

III. RESULTS AND DISCUSSION
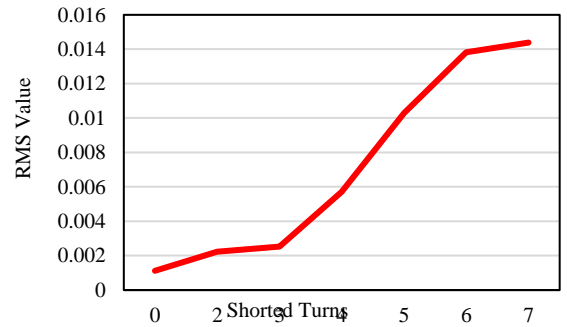
Fig. 3 RMS value of flux for different fault conditions

In the experiment model, two to seven turns short-circuit faults were made. Flux signal was obtained to analyze the

time domain waveform and EMD. The duration of the fault is one second. When the motor has a short-turn inter-turn short circuit, the stator flux amplitude and phase change are small. Fig. 2 show the experiment results in time domain waveform when there two to seven turns' short-circuit faults. These figures show that when there is a fault in the generator the amplitude of the flux rises. This behaviour is more obvious as the level of the fault increases. RMS values for normal and ITSCF of the stator winding is given in TABLE I. Result show that in normal condition the RMS value of flux in low, but in faulty condition, the flux signal have the higher values and this value increases as the level of the fault increases, which is shown in Fig. 3. The increment in RMS values show the fault severity level.

The EMD analysis of the stator flux signal is performed and the results are shown in Fig. 4 to Fig. 9. The EMD of the flux signal in the stator winding ITSCF yields 9 layers of empirical modal components IMFs. According to EMD analysis, when ITSCF occurs in the stator winding of the generator, the flux signal of the fault phase is decomposed by EMD, and the amplitudes of the IMF5 and IMF6 components will suddenly change at the time of fault. This is because, EMD decomposes the original signal into components of different frequency ranges, and a short-circuit fault in the stator winding will cause a new inter-turn flux to appear in the short-circuit loop. The envelope of the signal in these frequency ranges changes, so that the signal envelope contains the fault information. The sudden change of the IMF5 and IMF6 components at the time of the fault is caused by the fault of the generator. Therefore, a small short-circuit fault will not cause a significant change in the stator flux signal, but the fault of the generator can be detected by EMD, which shows its effectiveness. The peak to peak values of IMF5 and IMF6 increases with fault severity, which is shown in TABLE II and Fig. 10.
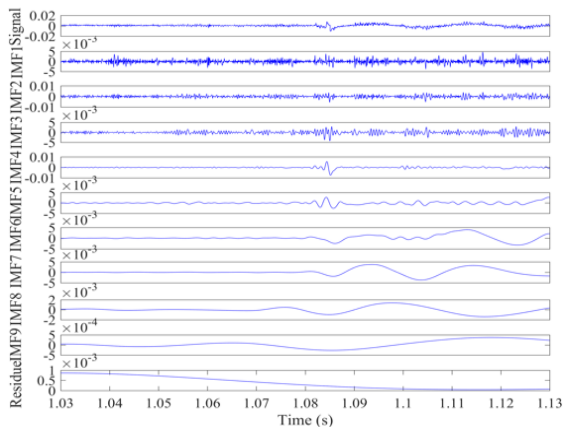
Fig. 4   EMD analysis of 2 turns short fault
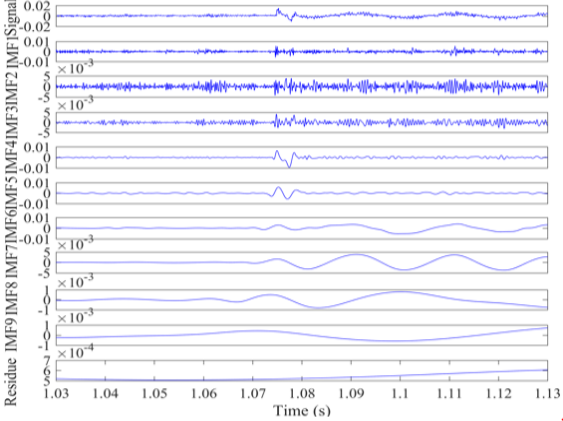
Fig. 5   EMD analysis of 3 turns short fault
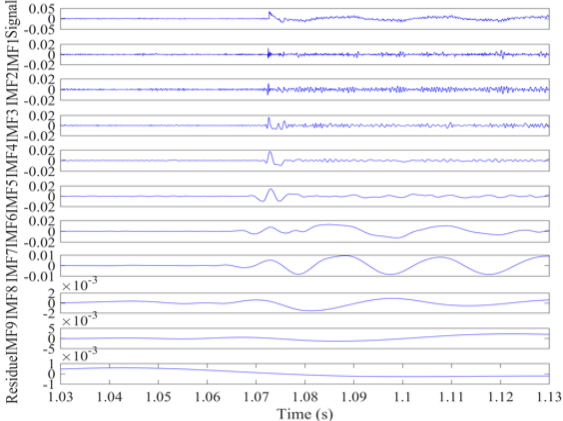
Fig. 6   EMD analysis of 4 turns short fault
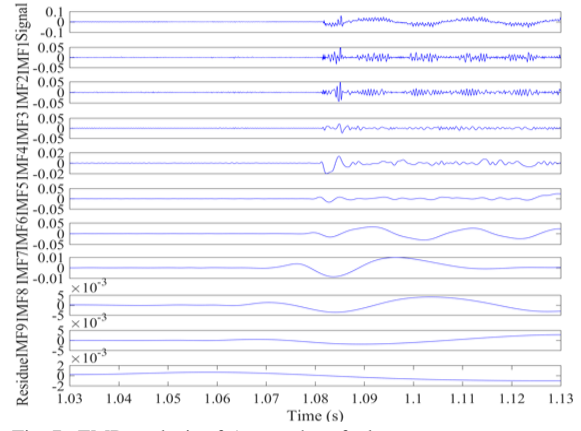
Fig. 7   EMD analysis of 5 turns short fault
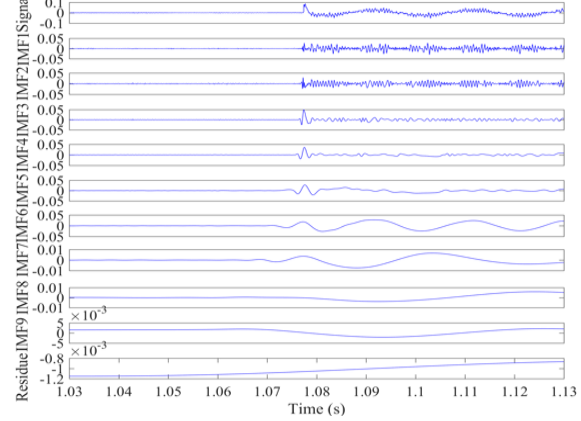
Fig. 8   EMD analysis of 6 turns short fault
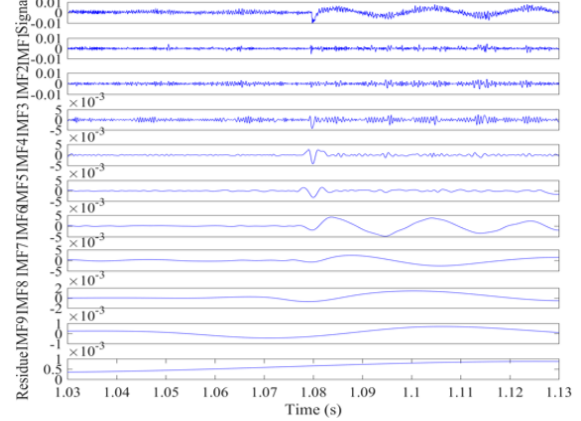
Fig. 9   EMD analysis of 7 turns short fault

TABLE II. PEAK TO PEAK VALUES OF IMF5 AND IMF 6 AT DIFFERENT FAULT CONDITIONS

Shorted turns	IMF 5	IMF 6
0	0.00084239	0.00078767
2	0.0053869	0.0071714
3	0.011345	0.0091967
4	0.024254	0.024893
5	0.041355	0.053644
6	0.048136	0.060104
7	0.057248	0.061497

IV. CONCLUSIONS
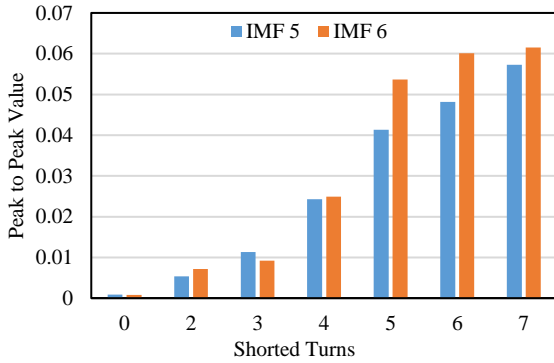
Fig. 10 Peak to peak values of IMF 5 and IMF 6

This research article is based on the experimental results which focus on inter turn short circuit fault detection of stator winding in doubly fed induction generator. Normal and two to seven turns short circuit fault was investigated at a full load condition. The magnetic flux antenna was placed on the axial position of the generator. Time domain and Empirical Mode Decomposition (EMD) analyses were found effective for the fault detection. The time domain analysis show that when the fault happened, the output voltage amplitude get higher. For EMD analysis of the flux in the stator winding ITSCF yields 9 layers of empirical modal components IMFs. The IMF5 and IMF6, was found useful to differentiate the normal and faulty condition of the machine. During the healthy condition, there is no sudden change in the amplitude of IMF5 and IMF6. But when the fault occurs, there was a sudden change in the amplitudes of the IMF5 and IMF6. From the results investigated in this research article concludes that for stator ITSCF detection in DFIG the time domain and EMD analyses on flux are efficacious than the other fault detection methods.

REFERENCES

[1] W. Qiao, and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis—Part I: Components and subsystems", IEEE Trans. Ind. Electron., vol. 62, no. 10, pp.6536-6545, 2015.

[2] A. S. Yunus, M. A. Masoum, and A. Abu-Siada, "Application of SMES to enhance the dynamic performance of DFIG during voltage sag and swell", IEEE Trans. on Applied Superconductivity, vol. 22 no. 4, pp. 5702009-5702009, 2012.

[3] M. Malekpour, B. T. Phung and E. Ambikairajah, "Modelling and diagnostic of incipient stator inter-turn short circuit fault in induction motors," IEEE Conf. Cond. Mon. and Diagn (CMD), pp. 1-6, Perth, 2018.

[4] C. Gerada et al., "The results do mesh," IEEE Ind. Appl. Mag., vol. 13, no. 2, pp. 62–72, Mar. 2007.

[5] Z. Hameed, S. H. Ahn, and Y. M. Cho, "Practical aspects of a condition monitoring system for a wind turbine with emphasis on its design, system architecture, testing and installation", Renewable Energy, vol. 35, no. 5, pp. 879-894, 2010.

[6] G. M. J Herbert, S. Iniyan, and D. Amutha, "A review of technical issues on the development of wind farms", Renewable & Sustainable Energy Reviews, vol. 32, pp. 619–41, 2014.

[7] J. Zhu, J. M. Yoon, D. He, Y. Qu, and E. Bechhoefer, "Lubrication oil condition monitoring and remaining useful life prediction with particle filtering", International Journal of Prognostics and Health Management, 4 (Special Issue Wind Turbine PHM), pp. 1-15, 2013.

[8] F. P. G. Márquez, A. M. Tobias, J. M. P. Pérez, and M. Papaelias, "Condition monitoring of wind turbines: Techniques and methods", Renewable Energy, vol. 46, pp. 169-178, 2012.

[9] N. Apergis, M. B. Jebli, and S. B. Youssef, "Does renewable energy consumption and health expenditures decrease carbon dioxide emissions? Evidence for sub-Saharan Africa countries", Rene Energy. vol. 127, pp. 1011–1016, 2018.

A KLIEP-based Transfer Learning Model for Gear Fault Diagnosis under Varying Working Conditions

Chao Chen, Fei Shen
School of Instrument Science and Engineering
Southeast University
Nanjing, China
chofy123@163.com
sfseu0311@163.com

Zhaoyan Fan
Department of Mechanical, Industrial and Manufacturing Engineering
Oregon State University
Cleveland, United States of America
Zhaoyan.Fan@oregonstate.edu

Robert X. Gao
Department of Mechanical and Aerospace Engineering
Case Western Reserve University
Cleveland, United States of America
Robert.Gao@case.edu

Ruqiang Yan*
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
yanruqiang@xjtu.edu.cn

Abstract—**Considering the fact that gear often works under varying working conditions, i.e., different domains, this paper presents a new approach that utilizes intrinsic time-scale decomposition (ITD) to extract features from vibration signals and transfer learning (TL) to achieve domain adaptation for gear fault diagnosis (GFD). The ITD first decomposes the vibration signal into several proper rotation components (PRCs). Then, the singular value vectors from the PRCs are extracted to provide differential indexes. TL aims to minimize the distribution distance among domains at most and solve the problem of distribution change, where a weight adjustment mechanism is involved in the Kullback-Leibler importance estimation procedure (KLIEP) for domain adaptation. Finally, the weighted domain vectors are applied to the GFD model. Experimental study performed on a drivetrain dynamics simulator (DDS) show that KLIEP performs well for domain adaption and the classification results also prove the superiority of proposed method in GFD when working condition changes.**

Keywords—Gearbox fault diagnosis, Transfer learning, KLIEP, varying working conditions.

I. INTRODUCTION

As one of the key parts of the rotating machinery, the running state of the gear has a great impact on the smooth operation of the system. In most of gear systems, the variation of interference noise and working conditions is unavoidable. Therefore, in order to maintain the long-term healthy operation of mechanical equipment, it is necessary to take the active gear fault diagnosis (GFD) strategies for preventing the occurrence of possible failures. However, varying working conditions faced in rotating machinery bring apparent consequence, such as: 1) accelerating the gear wear and shortening the gear life, and 2) adding extra noise on original vibration signals, thus increasing the difficulty of GFD [1-2].

To overcome the influence of varying working conditions, some feature extraction and pattern recognition algorithms have been applied to GFD of rotating machinery. For instance, Sharma [3] et al. reviewed GFD researches under various condition indicators; Cerrada [4] et al. proposed a new hierarchical feature selection model based on relative dependency for GFD; Cheng [5] built a planetary gear fault diagnosis model based on entropy feature fusion from signals' ensemble empirical mode decomposition. Zhang [6] et al. combined the grasshopper optimization algorithm (GOA) with variational mode decomposition (VMD) to build a gear diagnosis model when facing with variable conditions; Han [7] et al. proposed an enhanced convolutional neural network model to improve the fault diagnosis performance of planetary gearbox under variable working conditions. However, there are two limitations in most of current GFD models: 1) ever-changing conditions increase the complexity of proposed models, which tends to lack generalizability ; 2) different models need to be established to adapt to the change of working conditions, thus having high resource consumption. Inspired by transfer learning (TL) idea [8,9], a domain adaptation GFD model based on Kullback-Leibler importance estimation procedure (KLIEP) algorithm is presented in this paper to enhance the robustness of GFD model. This model is adopted by minimizing the feature distribution of KL distances of fault signals from different conditions to realize cross-domain learning, and to provide possible solutions for gear fault diagnosis under varying working conditions. The goal of transfer learning is to obtain effective gearbox diagnostic results when the machine runs in a new working environment while using the model built in the old working environment.

The rest of this paper is organized as follows. After theoretical background on transfer learning is introduced in section 2, the proposed GFD model based on KLIEP algorithm is presented in section 3. Then case studies for gearbox fault diagnosis will be conducted in section 4 to prove related performances. Finally, conclusions is summarized in section 5.

II. WORK BASIC METHODS

A. Transfer Learning

Compared to non-transfer learning strategies, three unique characteristics exist in the transfer learning models: multiple-tasks, cross-domains and different distribution. Its objective is to discover common knowledge hiding in source domains, and transfer to target domain [10,11]. Suppose there exist source domain $\mathcal{D}_s$, source task $\mathcal{T}_s$, target domain $\mathcal{D}_t$ and target task $\mathcal{T}_t$, the purpose of transfer learning is to

---

optimize predictive function f_t in target domain by incorporating the knowledge obtained from predictive function f_s in source domain. Here $\boldsymbol{D} = \{\mathcal{X}, \mathrm{P}(\boldsymbol{x})\}$ is represented by the feature space $\mathcal{X}$ and its marginal probability distribution (MPD), $\boldsymbol{x} = \{x_1, \cdots, x_n\} \in \mathcal{X}$. $\boldsymbol{T} = \{\mathcal{Y}, f\}$ is represented by the label space $\mathcal{Y}$, and the predictive function f. In non-transfer learning, $\boldsymbol{D}_s = \boldsymbol{D}_t$ and $\boldsymbol{T}_s = \boldsymbol{T}_t$, but in transfer learning, $\boldsymbol{D}_s \neq \boldsymbol{D}_t$ or $\boldsymbol{T}_s \neq \boldsymbol{T}_t$.

B. Influence of Working Conditions on GFD Signals

Generally, for gear fault diagnosis, the dynamics of gear components can be equivalent to the approximate second-order mass-spring-damping systems:

$$h(t) = \frac{1}{m\omega} e^{-\left(\xi/\sqrt{1-\xi^2}\right)\omega t} sin(\omega t) \tag{1}$$

where m represents the mass of vibratory system; ξ means the damping ratio; ω means the natural frequency. When the operating conditions change, the measured vibration signals will change. For a second-order mass-spring-damping system, the vibration characteristic depends on four parameters: mass, damping ratio, natural frequency and machine cycle. Different rotating speeds will change the machine cycle and bring the signal difference, causing differences of data distribution between GFD modeling and GFD application.

C. TL for GFD under Varying Working Conditions

Fig. 1 shows the transfer learning strategy for GFD under varying working conditions. It indicates that the training data in source working conditions are applied in the GFD model after transfer learning. Here, the data distribution between source working conditions and target working condition is different. Classical non-transfer strategy just utilizes the dataset under target working condition for both training and verification, while transfer learning strategy can build a more generic model by taking full advantage of the data in those source working conditions. The transfer learning can be successfully applied under two conditions: 1) the common knowledge exists in source working conditions and target working condition; 2) the amount of samples in target working condition is insufficient for individual GFD.
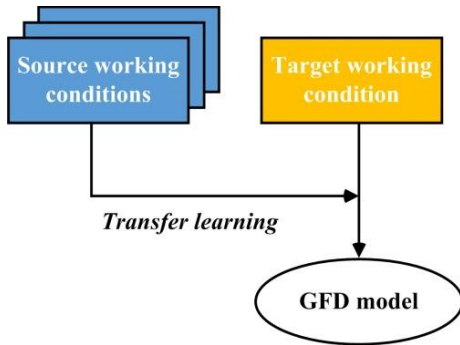
Fig.1. The transfer learning strategy for GFD under varying working conditions.

III. DOMAIN ADAPTATION BASED TL FOR GFD

A. Dataset and Feature Extraction

As a new time-frequency processing method, the intrinsic time-scale decomposition (ITD) [12,13] needs less calculation as compared to empirical mode decomposition,

thus having the advantages of fast speed. In ITD, the input vibration signal $\boldsymbol{X}_t$ can be adaptively decomposed into several proper rotation components (PRCs) and a residual term, shown as follows:

$$\boldsymbol{X}_t = \boldsymbol{L}\boldsymbol{X}_t + (1 - \boldsymbol{L})\boldsymbol{X}_t = \boldsymbol{L}_t + \boldsymbol{H}_t \tag{2}$$

where $\boldsymbol{L}$ represents the baseline extraction operator; $\boldsymbol{L}_t = \boldsymbol{L}\boldsymbol{X}_t$ represents the baseline signals; $\boldsymbol{H}_t = (1 - \boldsymbol{L})\boldsymbol{X}_t$ represents the separated PRCs.

Then the fully functional process of feature extraction process is listed as follows:

a) All extreme points of signals $\boldsymbol{X}_t$ and their corresponding time τ_k are calculated or obtained, and $\boldsymbol{M}$ represents the number of extreme points;

b) The piecewise linear baseline extraction operator $\boldsymbol{L}$ is defined on the interval of continuous extreme points $(\tau_k, \tau_{k+1}]$:

$$\boldsymbol{L}\boldsymbol{X}_t = \boldsymbol{L}_t = \boldsymbol{L}_k + \left(\frac{\boldsymbol{L}_{k+1} - \boldsymbol{L}_k}{\boldsymbol{L}_{k+2} - \boldsymbol{L}_k}\right)(\boldsymbol{X}_t - \boldsymbol{X}_k) \tag{3}$$

where,

$$\boldsymbol{L}_{k+1} = \alpha\left[\boldsymbol{X}_k + \left(\frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}\right)(\boldsymbol{X}_{k+2} - \boldsymbol{X}_k)\right] + (1 - \alpha)\boldsymbol{X}_{k+1} \tag{4}$$

where k=1, 2, …, $\boldsymbol{M}$-2; $0 < \alpha < 1$, (generally α=0.5).

c) The baseline signal $\boldsymbol{L}_t^1$ is separated from original signal $\boldsymbol{X}_t$, and the first PRC item is obtained:

$$\boldsymbol{H}_t^1 = (1 - \boldsymbol{L})\boldsymbol{X}_t = \boldsymbol{X}_t - \boldsymbol{L}_t^1 \tag{5}$$

Then, the $\boldsymbol{L}_t^1$ is considered as the original signal and repeat the step above until the baseline signal is a constant or monotone function. After that, original signal $\boldsymbol{X}_t$ can be decomposed as follows:

$$\begin{aligned}\boldsymbol{X}_t &= \left(\boldsymbol{H}\sum_{i=0}^{p-1}\boldsymbol{L}^i + \boldsymbol{L}^p\right)\boldsymbol{X}_t \\ &= \boldsymbol{H}_t^1 + \boldsymbol{H}_t^2 + \cdots + \boldsymbol{H}_t^p + \boldsymbol{L}_t^p\end{aligned} \tag{6}$$

where p represents the number of iterations; $\boldsymbol{H}_t^i$ represents the i-th inherent component and $\boldsymbol{L}_t^p$ represents the residual component.

d) Construct the matrix $\boldsymbol{C} = \left[\boldsymbol{H}_t^1; \boldsymbol{H}_t^2; \ldots; \boldsymbol{H}_t^p\right]$ and extract the singular values of matrix $\boldsymbol{C}$ as the feature vector of signals $\boldsymbol{X}_t$.

According to the steps above, the gear vibration signals collected from both source working condition and target working condition are shown as follows:

$$\boldsymbol{D}_s = \{(\boldsymbol{X}_s, \boldsymbol{Y}_s)\} = \left\{\left(\boldsymbol{x}_j^s, y_j^s\right)\right\}, j = 1,2,\ldots,N_s$$

$$\boldsymbol{D}_t = \{(\boldsymbol{X}_t, \boldsymbol{Y}_t)\} = \{(\boldsymbol{x}_i^t, y_i^t)\}, i = 1,2,\ldots,N_t$$

$$\boldsymbol{X} = \{\boldsymbol{X}_s, \boldsymbol{X}_t\}, \boldsymbol{Y}_s \in \boldsymbol{Y}, \boldsymbol{Y}_t \in \boldsymbol{Y} \tag{7}$$

where $\boldsymbol{x}_j^s$ and $\boldsymbol{x}_i^t$ represent the feature vectors; y_j^s and y_i^t represent the fault categories; N_s and N_t represent the sample number of source and target domains respectively.

As shown in section 2.2, signals under different working conditions bring differences of data distribution between $\boldsymbol{D}_s$ and $\boldsymbol{D}_t$. According to literature [14, 15], the

covariancedeviation in $\boldsymbol{D}_s$ can be corrected by weighting to achieve a similar target condition learning function. Here the weight $\beta = P_t(\boldsymbol{X},\boldsymbol{Y})/P_s(\boldsymbol{X},\boldsymbol{Y})$ can be simplified as $\beta = P_t(\boldsymbol{X})/P_s(\boldsymbol{X})$. Therefore, the Kullback-Leibler importance estimation procedure (KLIEP) algorithm can be used for gear fault diagnosis.

B. GFD Model based on KLIEP Algorithm

The basic principle of KLIEP model is to design an importance estimate β such that the Kullback-Leibler (KL) divergence between the weighted $\boldsymbol{D}_s$ and $\boldsymbol{D}_t$ is minimized. If two kinds of distributions are exactly the same, KL divergence distance will equal to zero. Therefore, the differences of the same fault type from both source domain and target domain needs to be reduced. Fig. 2 gives the specific flow diagram of GFD model with the KLIEP algorithm. Referring to [16], the weight β is assumed to be the linear model, which can be expressed as:

$$\beta(\boldsymbol{x}) = \sum_{l=1}^{N_t} \alpha_l K(\boldsymbol{x}, \boldsymbol{c}_l) \tag{8}$$

where $\alpha_l \geq 0$ is the KLIEP parameter that needs to be learned.

Due to the fact that $\boldsymbol{D}_s$ samples are easier to be obtained than $\boldsymbol{D}_t$ samples in GFD application, we assume that $N_t < N_s$. Here, $K(\cdot)$ represents the gauss kernel function and $\boldsymbol{c}_i$ is the center of gauss kernel function. Therefore, the KL divergence between $\boldsymbol{D}_s$ and $\boldsymbol{D}_t$ can be represented as:

$$\mathrm{KL}\big(P_t(\boldsymbol{x})\|\beta(\boldsymbol{x}) * P_s(\boldsymbol{x})\big)$$
$$= \sum P_t(\boldsymbol{x})\log(P_t(\boldsymbol{x})/\beta(\boldsymbol{x}) * P_s(\boldsymbol{x}))$$
$$= \sum P_t(\boldsymbol{x})\log\frac{P_t(\boldsymbol{x})}{P_s(\boldsymbol{x})} - \sum P_t(\boldsymbol{x})\log\beta(\boldsymbol{x}) \tag{9}$$

where the first item $\sum P_t(\boldsymbol{x})\log[P_t(\boldsymbol{x})/P_s(\boldsymbol{x})]$ is independent of actual weights, and the second item $\sum P_t(\boldsymbol{x})\log\beta(\boldsymbol{x})$ can be further calculated as:

$$\sum P_t(\boldsymbol{x})\log\beta(\boldsymbol{x}) \approx \sum_{i=l}^{N_t} \log\beta(\boldsymbol{x}_i^t)/N_t$$
$$= \sum_{i=l}^{N_t} \log\sum_{l=1}^{N_t} \alpha_l K(\boldsymbol{x}, \boldsymbol{c}_l)/N_t \tag{10}$$

Besides,

$$\sum \beta(\boldsymbol{x})P_t(\boldsymbol{x}) \approx \sum_{j=l}^{N_s} \log\beta(\boldsymbol{x}_i^t)/N_s$$
$$= \sum_{j=l}^{N_s} \log\sum_{l=1}^{N_t} \alpha_l K(\boldsymbol{x}, \boldsymbol{c}_l)/N_s \tag{11}$$

Finally, the minimum KL divergence in equation (9) could be converted to a new convex optimization problem:

$$\max\left[\sum_{i=l}^{N_t} \log\sum_{l=1}^{N_t} \alpha_l K(\boldsymbol{x}_i^t, \boldsymbol{c}_l)\right]$$
$$\text{s.t. } \sum_{j=l}^{N_s} \log\sum_{l=1}^{N_t} \alpha_l K(\boldsymbol{x}_j^s, \boldsymbol{c}_l)/N_s = 1, \alpha_l \geq 0 \tag{12}$$

Therefore, the specific procedure for GFD model with the KLIEP algorithm can be listed as:

a) Allocating the $\boldsymbol{D}_s$ and $\boldsymbol{D}_t$ samples, respectively;

b) Initializing parameter $\alpha_l = [\alpha_1, \alpha_2, \ldots, \alpha_{N_t}]$;

c) Iteration: $ite = 1,2,\ldots$:

c-1, Choosing a subset from $\boldsymbol{D}_s$ as the data centers of gauss function randomly;

c-2, Computing the global optimal solution $\alpha_1, \alpha_2, \ldots, \alpha_{N_t}$ using gradient ascent method and k-fold cross validation;

c-3, If the KL divergence in current iteration is bigger than last iteration, the GFD mode will be output, otherwise the iteration will be continued;

d) Computing the final weight $\beta(\boldsymbol{X}_s) = \sum_{l=1}^{N_t} \alpha_l K(\boldsymbol{X}_s, \boldsymbol{c}_l)$ based on the α_l value, and building a target-oriented GFD model using the weighted source domain data.
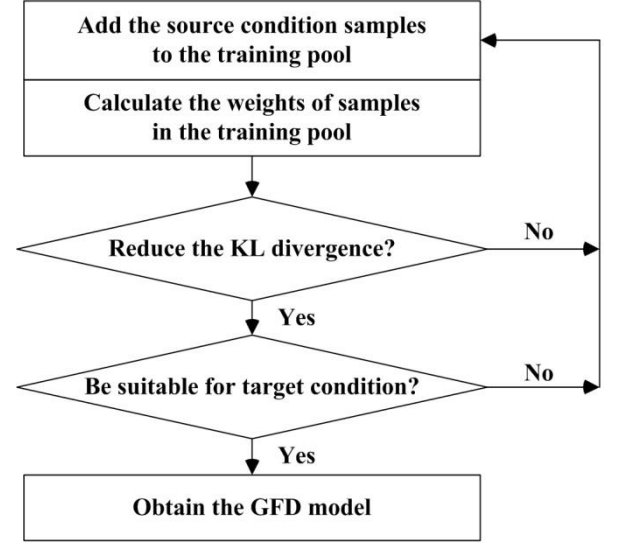
Fig.2. The specific flow diagram of GFD model with the KLIEP algorithm.

IV. EXPERIMENT AND ANALYSIS

A. Test Rig and Data Description

Experimental study was performed on a drivetrain dynamics simulator (DDS) in Yan's group [17], which is shown in Fig. 3, and vibration signals were measured from six channels installed on reduction gearbox and planetary gearbox. Testing gears contain five categories: normal condition (NC), root crack fault (RCF), chipped tooth fault (CTF), miss tooth fault (MTF), surface wear fault (SWF). Rotating speeds were set from 20 to 50Hz. Loads were set from 0 to 12.8N·m [17]. The type of gear faults and its experimental conditions are shown in Table 1.
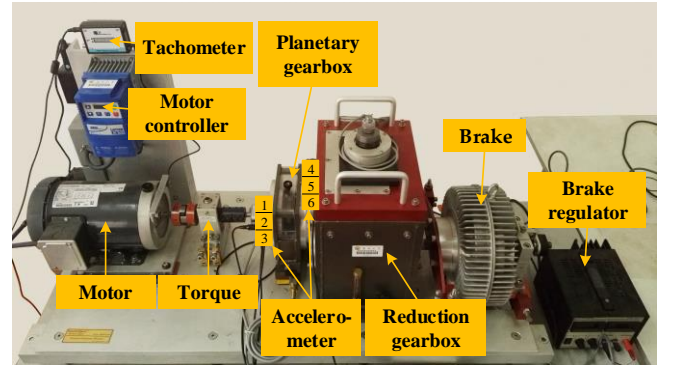
Fig.3. Spectra Quest's Drivetrain Dynamics Simulator.

Faults	C1 NC	C2 RCF	C3 SWF	C4 MTF	C5 CTF
Speeds (Hz)	D1 20	D2 30	D3 40	D4 50	—
Load (N*m)	E1 0	E2 3.657	E3 7.315	E4 10.97	E5 14.63
Location	F1 Planetary gearbox		F2 Reduction gearbox		—

B. Domain Adaptation Experiments

During the domain adaptation experiments, the target working condition is considered as: T={D3, E1, F2}, and the source working conditions are considered as: $S1$={D1, E1, F2}; $S2$={D2, E1, F2}; $S3$={D3, E2, F2}; $S4$={D4, E1, F2}; $S5$={D5, E1, F2}; $S6$={D3, E3, F2}; $S7$={D1, E1, F2}. It is indicated that the each dataset S changes one of the 3 working conditions of dataset T.

Firstly, to verify the effectiveness of feature extraction and to analyze model's clustering performance, the visualization tool t-SNE [18] is used, which aims to map origin high-dimension features to two-dimension features, as shown in Fig. 4. Notice that, the clustering performance of original vibration data is poor, thus being difficult to distinguish. On the contrary, the t-SNE graph extracted by ITD model shows strong clustering performance, which is suitable for identification of gear fault signals.

(a) Original features

(b) ITD features
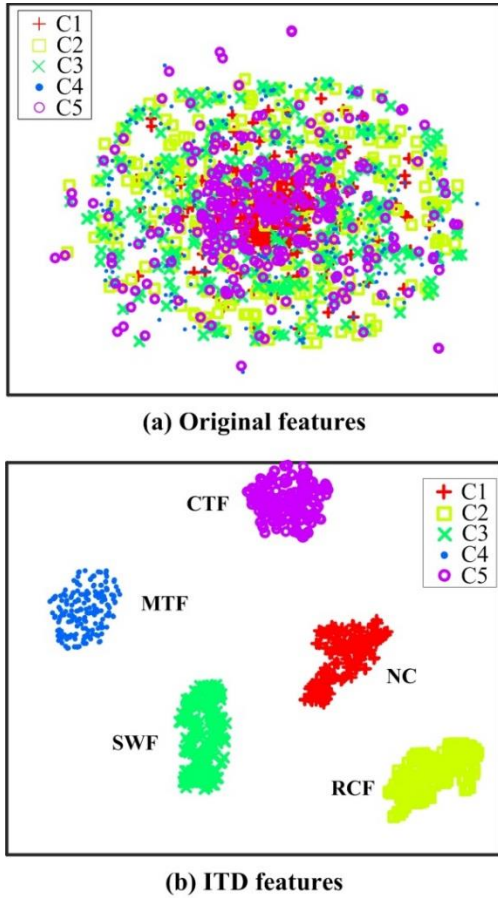
Fig.4. The visualized t-SNE graphs.

Then, Fig. 5 compares the KL divergences between seven kinds of source working conditions from $S1$ to $S7$. It can be seen that the influence of working conditions on the

GFD results is ordered by: rotating speed < load < location. In the KLIEP algorithm, the influence of rotating speed can be weakened, thus being helpful to industrial GFD application. Meanwhile, the difference between source domain and target domain can be weakened using the KLIEP algorithm, by comparing with original KL divergences.

C. Gear Fault Diagnosis Experiments

The GFD experiments include variable rotating speed transfer learning (from $S5$ to T) and variable load transfer learning (from $S3$ to T). In this part, some methods are compared as follows: Support Vector Machine (SVM), Factor Analysis (FA), Maximum Mean Discrepancy (MMD), Deep CORAL (FA) and Domain Adversarial Neural Network (DANN, also known as RevGrad). In those aforesaid models, the source domains are only used for training.
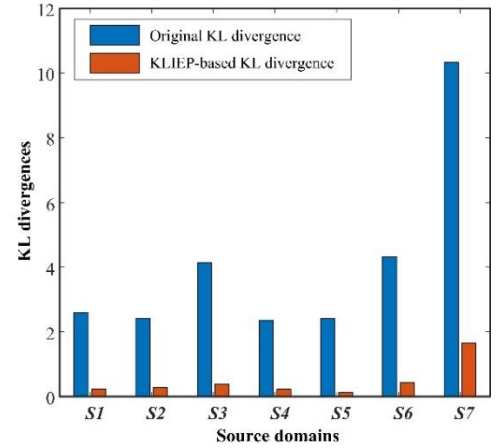
Fig.5. The KL divergences between seven kinds of source working conditions.

(a) Transfer from $S5$ to T

(b) Transfer from $S3$ to T
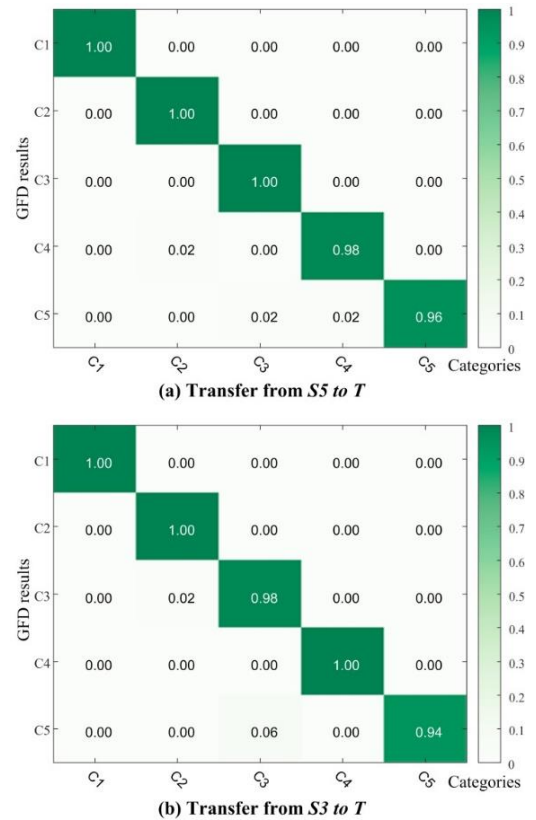
Fig.6. Confusion matrix of GFD classification results with the KLIEP model.

TABLE II. THE MAE AND RMSE INDEXES IN [*S5*, *T*]

Fault type / Model	NC MAE/RMSE	RCF MAE/RMSE	MTF MAE/RMSE	SWF MAE/RMSE	CTF MAE/RMSE
SVM	0/0	0.02/0.14	0.52/0.96	0.06/0.25	0.6/1.19
FA-SVM	0/0	0.06/0.43	0.52/0.97	0.02/0.14	0.36/0.85
MMD-SVM	0.22/0.71	0.16/0.49	0.14/0.47	0.04/0.2	0.08/0.35
RevGrad	0/0	0/0	0.02/0.14	0/0	0.12/0.49
D-Coral	0/0	0/0	0/0	0.04/0.28	0.06/0.32
KLIEP-SVM	0/0	0.02/0.14	0/0	0/0	0.08/0.45

TABLE III. COMPARISON BETWEEN DIAGNOSTIC ACCURACIES OF DIFFERENT GFD MODELS

Model / Data set	SVM	FA-SVM	MMD-SVM	RevGrad	D-Coral	KLIEP-SVM
[*S5*, *T*]	0.856	0.878	0.924	0.986	**0.988**	**0.988**
[*S3*, *T*]	0.812	0.846	0.918	0.944	0.946	**0.984**

Fig. 6 gives the confusion matrix of GFD classification results with the KLIEP model. From this figure, two main conclusions can be obtained: 1) by comparing the diagnostic results between varying rotating speed and load, the GFD performances are close (98.8% & 98.4%). Although the influence of varying loads is larger than varying rotating speeds, the KLIEP model can effectively reduce the weakness of the former and get a satisfying result; 2) under the same working condition, the recognition accuracy (96.0% & 94.0%) of CTF type is lower than other fault categories. Although the sensitivity of KLIEP to CTF type is relatively low, it is still acceptable. Table II gives the MAE and RMSE indexes in [*S5*, *T*] and Table III lists the total GFD accuracy of [*S5*, *T*] and [*S3*, *T*]. For a fair comparison, the radial basis function (RBF) is chosen as kernel function in all kernel-based methods, where grid searching and 5-fold cross validation are utilized to seek optimal penalty and kernel function parameters. In order to quantify the GFD performance, the mean absolute error (MAE) and root mean squared error (RMSE) are calculated:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\overset{\circ}{y}_i - y_i\right|$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\overset{\circ}{y}_i - y_i\right)^2} \qquad (13)$$

Two conclusions can be obtained from Table II and III: 1) although both varying rotating speed and load may change the feature distribution of vibration signal, they have little impact on the adaptive model based on KLIEP algorithm. Therefore, the proposed method is suitable for GFD under variable working conditions; 2) KLIEP model has the highest diagnostic performance among five models, and it can improve the diagnosis accuracy by up to 17.7% compared with non-transfer models such as SVM. The results in Table III prove that the proposed method improves the reliability of gear fault diagnosis under variable working conditions.

V. CONCLUSION

A domain adaptation model for gear fault diagnosis under different working conditions is proposed. 1) As a cross-domain learning method, KLIEP algorithm weakens the sample distribution difference between source and target working conditions by weighting the data of the former. This algorithm overcomes the limitation of the same distribution of training and test data required by traditional machine learning, and provides a valuable idea for the engineering application of GFD. 2) By comparing the KL distances between the data of source and target working conditions, it was found that the influence of varying rotating speeds was slightly lower than varying loads. 3) By compared with traditional methods, the diagnostic accuracy of the adaptive model is the highest, with an average improvement of 11%~17%.

REFERENCES

[1] F. H. Li, R. Li, L. Tian, L. Chen, J. Liu, "Data-driven time-frequency analysis method based on variational mode decomposition and its application to gear fault diagnosis in variable working conditions," Mechanical Systems and Signal Processing, vol. 116, pp. 462-479, 2019.

[2] Y. B. Li, K. Feng, X. Liang, M. J. Zuo, "A fault diagnosis method for planetary gearboxes under non-stationary working conditions using improved Vold-Kalman filter and multi-scale sample entropy," Journal of Sound and Vibration, vol. 439, pp. 271-286, 2019.

[3] V. Sharma, A. Parey, "A Review of Gear Fault Diagnosis Using Various Condition Indicators," Procedia Engineering, vol. 144, pp. 253–263, 2016.

[4] M. Cerrada, R. V. Sánchez, F. Pacheco, D. Cabrera, G. Zurita, C. Li, "Hierarchical feature selection based on relative dependency for gear fault diagnosis," Applied Intelligence, vol. 44, no. 3, pp. 687–703, 2016.

[5] G. Cheng, X. H. Chen, H. Y. Li, P. Li, H. G. Liu, "Study on planetary gear fault diagnosis based on entropy feature fusion of ensemble empirical mode decomposition," Measurement, vol. 91, pp. 140–154, 2016.

[6] Z. Xin, M. Qiang, Z. Heng, W. Lei, "A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery," Mechanical Systems and Signal Processing, vol. 108, pp. 58-72, 2018.

[7] Y. Han, B. P. Tang, L. Deng, "An enhanced convolutional neural network with enlarged receptive fields for fault diagnosis of planetary gearboxes," Computers in Industry, vol. 107, pp. 50-58, 2019.

[8] M. H. Qiu, L. Yang, F. Ji, W. Zhou, J. Huang, H. Q. Chen, W. B. Croft, W. Lin, "Transfer Learning for Context-Aware Question Matching in Information-seeking Conversations in E-commerce," Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, vol. 2, pp. 208–213, 2018.

[9] K. Weiss, T. M. Khoshgoftaar, D. D. Wang, "A survey of transfer learning," Journal of Big Data, vol. 3, no. 1, pp. 1–40, 2016.

[10] B. Zoph, D. Yuret, J. May, K. Knight, "Transfer Learning for Low-Resource Neural Machine Translation," Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 1568–1575, 2016.

[11] Q. Wang, G. Michau, O. Fink, "Domain adaptive transfer learning for fault diagnosis," 2019 Prognostics and System Health Management Conference (PHM-Paris), pp. 279–285, 2019.

[12] Z. Q. Xing, J. F. Qu, Y. Chai, Q. Tang, Y. M. Zhou, "Gear fault diagnosis under variable conditions with intrinsic time-scale decomposition-singular value decomposition and support vector machine," Journal of Mechanical Science and Technology, vol. 31, no. 2, pp. 545-553, 2017.

[13] J. Yu, H. Liu, "Sparse coding shrinkage in intrinsic time-scale decomposition for weak fault feature extraction of bearings," IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 7, pp. 1579-1592, 2018.

[14] J. J. Li, K. Lu, Z. Huang, L. Zhu, H. T. Shen, "Transfer independently together: a generalized framework for domain adaptation," IEEE transactions on cybernetics, vol. 49, no. 6, pp. 2144-2155, 2018.

[15] S. J. Huang, J. W. Zhao, Z. Y. Liu, "Cost-effective training of deep cnns with active model adaptation," Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp.19-23, 2018.

[16] K. Y. Yan, W. M. Zheng, Z. Cui, Y. Zong, T. Zhang, C. G. Tang, "Unsupervised facial expression recognition using domain adaptation based dictionary learning approach," Neurocomputing, vol. 319, pp. 84-91, 2018.

[17] R. Q. Yan, F. Shen, C. Sun, X. F. Chen, "Knowledge Transfer for Rotary Machine Fault Diagnosis," IEEE Sensors Journal, vol. 20, no. 15, pp. 8374-8393, 2020.

[18] N. Pezzotti, B. P. F. Lelieveldt, L. v. d. Maaten, T. Höllt, E. Eisemann, A. Vilanova, "Approximated and User Steerable tSNE for Progressive Visual Analytics," IEEE Transactions on Visualization and Computer Graphics, vol. 23, no. 7, pp. 1739-1752, 2017.

Battery Evaluation Based on Mechanism Parameters

Xiaofang Cheng*
dept.of Thermal Science and Energy Engineering
University of Science and Technology of China
Hefei, China
xfcheng@ustc.edu.cn

Mao Lei
dept.of Precis Machinery & Precis Instrumentat
University of Science and Technology of China
Hefei, China
leimao82@ustc.edu.cn

Li Wen
dept.of Thermal Science and Energy Engineering
University of Science and Technology of China
Hefei, China
liwen96@mail.ustc.edu.cn

Zhang Chen
dept.of Precis Machinery & Precis Instrumentat
University of Science and Technology of China
Hefei, China
zhang985@mail.ustc.edu.cn

Abstract—**The maximum power is the most direct index to evaluate the quality of battery. In order to find the maximum power, experiments can be tried again and again; as for the battery model, to establish it in advance, then to obtain the derivative. By using the relationship between function expression and function derivation under maximum power, the battery can be evaluated by the first derivative value of volt ampere function at the maximum power. Further investigation of the symbol state of the second derivative of the volt ampere function can evaluate the battery performance under the single rational function without paying attention to the maximum power condition. Based on the function of the first and second derivative of the volt ampere function in the battery evaluation, a method of battery evaluation, rooted in the battery parameters in the volt ampere function, is constructed. Via the method, the derivation and deduction of the maximum power in the model can be avoided, and the repeated search for the maximum power in the experiment can also be avoided.**

Keywords—battery potential parameters, battery mechanism parameters, maximum power, battery evaluation

I. BACKGROUND AND PROBLEMS

In recent decades, the research on battery is in full swing, and the evaluation of battery quality has emerged as the times require. [1,2]

The maximum power is the most direct index to evaluate the battery. Battery power can be written as,

$$p_m = I_m V_m = \begin{cases} \xrightarrow{\text{Voltage source battery}} \dfrac{V_m^2}{R_{mL}} \\ \xrightarrow{\text{Current source battery}} I_m^2 R_{mL} \end{cases} \quad (1)$$

$$\xrightarrow{\text{Mathematical expression}} p(x_m) = x_m f(x_m) = x_m^2 \Omega_{mL}$$

According to (1), the value Ω_{mL} can be used as the evaluation index of battery power. More higher the value of index, more better the battery power.

Therefore, the relationship between Ω_{mL} and $f(x_m)$ needs establishing. For linear volt ampere batteries, this relationship has been established; for voltage source batteries, the value equals that of the internal resistance of series batteries $(\Omega_{mL} = R_s)$; for current source batteries,

the value is equal to that of the internal resistance of batteries in parallel $(\Omega_{mL} = R_s)$, but ,as for the battery with nonlinear volt ampere, this relationship is not established, yet[3].

II. BATTERY POWER

The product of current and voltage is called as battery power, and the mathematical expression of the battery power, and the maximum power are

$$p(x) = xf(x)$$

$$\xrightarrow{\text{derivative}} p'(x) = f(x) + xf'(x) \quad (2)$$

$$\xrightarrow[\text{maximum power}]{p'(x)|_{x_m}=0} f(x_m) + x_m f'(x)\big|_{x_m} = 0$$

Where: $p(x)$ is the electric power; $f(x)$ is the volt ampere function, expressed as voltage or current; x is the independent variable, shown as the current or voltage; x_m is the independent variable value at the maximum power.

Therefore, the maximum power of the battery can be written in two ways.

$$p(x_m) = x_m f(x_m) = -x_m^2 f'(x)\big|_{x_m} \quad (3)$$

The conclusion is drawn by comparing with the two formulas: (1) and (3): $\Omega_{mL} = -f'(x)\big|_{x_m}$.

Comparison on the maximum power of the two batteries (i, j) Supposing that the independent variables of the two batteries are equal under the maximum power condition, the comparison of the batteries performance would be converted into the numerical comparison of the first derivative of the volt ampere function at the maximum power.

$$\frac{p_i(x_{m_i})}{p_j(x_{m_j})} = \frac{x_{m_i}^2 f_i'(x)\big|_{x_{m_i}}}{x_{m_j}^2 f_j'(x)\big|_{x_{m_j}}} \xrightarrow[\text{Comparison condition}]{x_{m_i}=x_{m_j}=x_m} \frac{p_i(x_m)}{p_j(x_m)} = \frac{f_i'(x)\big|_{x_m}}{f_j'(x)\big|_{x_m}} \quad (4)$$

III. Battery Volt Ampere

The volt ampere curve of battery can be expressed as the following general function.[4]

$$f(x) = A - \sum_{k=1}^{3} f_k(x) \tag{5}$$
$$= A - Bx - C\ln(C_1 x + C_2) - D[\exp(D_1 x) - D_2]$$

Where: A is the battery potential parameter, which is given by battery principle; the other three are mechanism functions, (B, C, D) are the mechanism parameters of battery's dependent variable; (C_1, D_1) are the mechanism parameters of battery's independent variable; (C_2, D_2) are the parameters of battery experiment.

Each mechanism function of battery is independent mutually, that is, each mechanism function has its own law, but not be affected by other mechanism functions. According to this principle, $(x = 0)$ is put into the general function of battery volt ampere (4), and noted that $f(0) = A$.

$$f(0) = A - \sum_{k=1}^{3} f_k(0)$$
$$= A - C\ln(C_2) - D[1 - D_2] = A$$
$$\Rightarrow -C\ln(C_2) - D[1 - D_2] = 0 \tag{6}$$

$$\xrightarrow[\text{correction}]{\text{The mutually independent mechanism functions}} \begin{cases} C_2 = 1 \\ D_2 = 1 \end{cases}$$

It can be deduced that the value of battery's experimental parameters (C_2, D_2) must be equal to "1", otherwise, the independence of the mechanism will be destroyed.

So the correct form o battery's volt ampere function is

$$f(x) = A - \sum_{k=1}^{3} f_k(x)$$
$$= A - Bx - C\ln(C_1 x + 1) - D[\exp(D_1 x) - 1]$$
$$= A - \sum_{k=1}^{3} \begin{cases} \xrightarrow[\text{Mechanism}]{\text{Linear}} f_1(x) = f_B(x) = Bx \\ \xrightarrow[\text{Mechanism}]{\text{Nonlinear}} \begin{cases} \rightarrow f_2(x) = f_C(x) = C\ln(C_1 x + 1) \\ \rightarrow f_3(x) = f_D(x) = D[\exp(D_1 x) - 1] \end{cases} \end{cases} \tag{7}$$

And at $(x = x_{\max}, x = 0)$, there are

$$\begin{cases} f(x_{\max}) \\ = A - Bx_{\max} - C\ln(C_1 x_{\max} + 1) - D[\exp(D_1 x_{\max}) - 1] = 0 \\ f(x \rightarrow 0) = A \end{cases} \tag{8}$$

Then： $x_{\max} = f^{-1}(A, B, C, C_1, D, D_1)$ and $(x = 0) = f^{-1}(A)$.

IV. Maximum Battery Power Expressed By Volt Ampere Derivative

By derivative calculation of the volt ampere function (7), the potential parameters of battery, A can be

eliminated and the mechanism parameters of battery can be left:

$$\begin{cases} f'(x) \\ = -[B + CC_1(C_1 x + 1)^{-1} + DD_1 \exp(D_1 x)] = -\sum_{k=1}^{3} f_k'(x) \\ f''(x) \\ = -[-CC_1^2 (C_1 x + 1)^{-2} + DD_1^2 \exp(D_1 x)] = -\sum_{k=1}^{3} f_k''(x) \end{cases} \tag{9}$$

That is, the derivative of volt ampere function is equal to the sum of the derivatives of mechanism function, the difference between them is the sign '-'.

(7) and (9) are put into (3), two expressions of maximum battery power are obtained

$$p(x_m) = x_m f(x_m)$$
$$= x_m A - Bx_m - C\ln(C_1 x_m + 1) - D[\exp(D_1 x_m) - 1] \tag{10}$$
$$p(x_m) = x_m^2 \times \sum_{k=1}^{3} f_k'(x)\big|_{x_m}$$
$$= x_m^2 [B + CC_1(C_1 x_m + 1)^{-1} + DD_1 \exp(D_1 x_m)]$$

The formula (10) reveals that the maximum power of battery is the function constituted of potential parameters and mechanism parameters.

Table 1 shows the expression of the maximum power of battery under each single mechanism.

V. Battery Evaluation Based On Mechanism Parameters

The maximum power of battery parameter type is regarded as the summation of mechanism function. The value of independent variable $(x_m \rightarrow 0)$ under maximum power can be rewritten and put into equation (11).

$$\bar{p}(x_m \rightarrow 0)$$
$$= \sum_{k=1}^{3} f_k'(x)\big|_{x_m \rightarrow 0} \tag{11}$$
$$= B + CC_1 + DD_1$$
$$\xrightarrow{C=D=0} \bar{p}(x_m \rightarrow 0) = B$$

Then only the mechanism parameters of battery (B, C, C_1, D, D_1) are included in the evaluation formula.

If not considering the nonlinear mechanism parameters (C, C_1, D, D_1), only the linear mechanism parameter B (i.e. the internal resistance of the battery) is left. There are actually such battery evaluation techniques.

The existence of nonlinear mechanism leads to the following conditions in the sum of second derivative of mechanism function $\sum_{k=1}^{3} f_k''(x)$.

TABLE 1   MAXIMUM POWER OF SINGLE MECHANISM

Name	Linear Mechanism	Logarithmic Mechanism	Exponential Mechanism
Volt ampere function $f(x)$	$A - Bx$	$A - C\ln(C_1 x + 1)$	$A - D\left[\exp(D_1 x) - 1\right]$
Mechanism function $f'_k(x)$	Bx	$C\ln(C_1 x + 1)$	$D\left[\exp(D_1 x) - 1\right]$
First Derivative of Mechanism $f'_k(x)$	B	$CC_1(C_1 x + 1)^{-1}$	$DD_1\exp(D_1 x)$
$f(x_m) + x_m f'(x)\big|_{x_m} = 0$	$2x_m = \dfrac{A}{B}$	$x_m C_1(C_1 x_m + 1)^{-1}$ $+\ln(C_1 x_m + 1) = \dfrac{A}{C}$	$(1 + x_m D_1)\exp(D_1 x_m)$ $= \dfrac{A+D}{D}$
$f(x_{\max}) = 0$	$x_{\max} = \dfrac{A}{B}$	$\ln(C_1 x_{\max} + 1) = \dfrac{A}{C}$	$\exp(D_1 x_{\max}) = \dfrac{A+D}{D}$
$p_k(x_m)$ $\dfrac{x_m f(x_m)}{x_m^2 f'_k(x)\big|_{x_m}}$	$\dfrac{Bx_m^2}{Bx_m^2}$	$\dfrac{x_m^2 \times CC_1(C_1 x_m + 1)^{-1}}{x_m^2 \times CC_1(C_1 x_m + 1)^{-1}}$	$\dfrac{x_m^2 \times DD_1\exp(D_1 x_m)}{x_m^2 \times DD_1\exp(D_1 x_m)}$

$$\sum_{k=1}^{3} f''_k(x) = DD_1^2\exp(D_1 x) - CC_1^2(C_1 x + 1)^{-2}$$

$$\xrightarrow{\;x=0\;} \sum_{k=1}^{3} f''_k(0) = DD_1^2 - CC_1^2 \qquad (12)$$

$$\Rightarrow \begin{cases} \xrightarrow{DD_1^2 > CC_1^2} \sum_{k=1}^{3} f''_k(x) > 0 \\[2mm] \xrightarrow{DD_1^2 = CC_1^2} \sum_{k=1}^{3} f''_k(x) = 0 \Leftrightarrow f''_B(x) = 0 \\[2mm] \xrightarrow{DD_1^2 < CC_1^2} \sum_{k=1}^{3} f''_k(x) < 0 \end{cases}$$

Therefore, It is necessary to re-establish the battery power evaluation under the mechanism parameters.

A. *Battery evaluation under single mechanism operation*

First of all, the evaluation is conducted in the batteries under each mechanism, and the comparisons on the performance of battery types under different mechanism are clarified. See Table 2. The sign state of the second derivative reflects the opening direction of the voltammetric curve. As for the power located in the first quadrant, the opening direction is downward, And the volt ampere area with the downward opening direction, is obviously larger than that with the upward opening direction. Table 1 shows the performance of single mechanism battery.

The second derivative value does not affect the evaluation of battery quality, so it can be applied $(x = 0)$. Now that the results of the independent variable value do not change the sign direction of the derivative, the value of the independent variable in the first derivative could also be ignored, but would not affect the results of battery evaluation.

B. *Battery evaluation under mechanism summation*

In order to keep all battery parameters and consider the influence of nonlinear mechanism parameters, the battery evaluation formula is designed as follows

$$E = E(x_m \to 0)$$

$$= \frac{\sum_{k=1}^{3} f''_k(x)\big|_{x_m \to 0}}{\sum_{k=1}^{3} f'_k(x)\big|_{x_m \to 0}} = \frac{DD_1^2 - CC_1^2}{B + CC_1 + DD_1} \qquad (13)$$

According to (13), the battery evaluation under various mechanisms, shown in Table 3 ,can be obtained, and the ranking of advantages and disadvantages is given in the last line.

Corresponding to the research and use of the actual battery, there are :

Mechanism Function	$f_k(x)$	$f'_k(x)$	$f''_k(x)$	Performance of battery
$f_C(x)$-2	$C\ln(C_1 x + 1)$	$CC_1(C_1 x + 1)^{-1}$	$-CC_1^2(C_1 x + 1)^{-2}$	inferior
$f_B(x)$-1	Bx	B	0	middle
$f_D(x)$-3	$D\left[\exp(D_1 x) - D_2\right]$	$DD_1\exp(D_1 x)$	$DD_1^2\exp(D_1 x)$	excellent

$$\begin{cases} 1. E_D^{\text{Ideal photovoltaic}} > E_{BD}^{\text{Actual photovoltaic}} \\ > E_{D>C} > \left(E^{\text{fuel cell}}, E_B^{\text{dry battery}}, E_{C=D} \right) \\ 2. > E_{C>D} > E_C > E_{BC} \end{cases} \quad (14)$$

It can be seen that, as for the batteries which are being studied and used, the maximum power evaluation reveals it lies in the advantageous range of choice in (14).

ACKNOWLEDGMENT

REFERENCES

[1] JinWeiren,PangJing,Tangling,Yuzheng, "Research Progress of Consistency Evaluation Method for lithium Ion Power Batteries," Bettery，44(1), 53-56,2014.

[2] Muzhong Shen, Yulong Ding and Keith Scott, "A Concise Model for Evaluating Water Electrolysis". International Journal of Hydrogen Sources 36(2011), 14335

[3] Lincheng,Mengxiangfeng,Wangzhenbo,Sunfengchun, "Research on Comprehensive Performance Evaluation Method of Power Battery for Electric Vehicle,"High technology communication,2006.pp.929-933

[4] James Lamir, Andrew Dicks, et al. Fuel cell system - Principle design application [M]. Beijing. Science Press, 2006.

TABLE 3 BATTERY EVALUATION UNDER VARIOUS MECHANISMS

Three mechanisms	$E = \dfrac{DD_1^2 - CC_1^2}{B + CC_1 + DD_1}$		
Single mechanism	$E_C = \dfrac{-CC_1^2}{CC_1} = -C_1$	$E_B = \dfrac{0}{B}$	$E_D = \dfrac{DD_1^2}{DD_1} = D_1$
Double mechanism	$E_{BC} = \dfrac{-CC_1^2}{B + CC_1}$	$E_{CD} = \dfrac{DD_1^2 - CC_1^2}{CC_1 + DD_1}$	$E_{BD} = \dfrac{DD_1^2}{B + DD_1}$
Single / Double mechanism	$\dfrac{E_C}{E_{BC}} = \dfrac{B + CC_1}{CC_1}$	$E_B E_{CD} = \dfrac{0 \times (DD_1^2 - CC_1^2)}{B(CC_1 + DD_1)}$	$\dfrac{E_D}{E_{BD}} = \dfrac{B + DD_1}{DD_1}$
Double / Three mechanism	$\dfrac{E}{E_{BC}} = \dfrac{(B + CC_1)(DD_1^2 - CC_1^2)}{CC_1^2(B + CC_1 + DD_1)}$	$\dfrac{E}{E_{CD}} = \dfrac{CC_1 + DD_1}{B + CC_1 + DD_1}$	$\dfrac{E}{E_{BD}} = \dfrac{(B + DD_1)(CC_1^2 - DD_1^2)}{DD_1^2(B + CC_1 + DD_1)}$
Single / Three mechanism	$\dfrac{E}{E_C} = \dfrac{CC_1^2 - DD_1^2}{C_1(B + CC_1 + DD_1)}$	$EE_B = \dfrac{0 \times (DD_1^2 - CC_1^2)}{B(B + CC_1 + DD_1)}$	$\dfrac{E}{E_D} = \dfrac{DD_1^2 - CC_1^2}{D_1(B + CC_1 + DD_1)}$
Ranking of advantages and disadvantages	$E_{BC} < E_C < E_{D<C} < E_B, E_{D=C}, E < E_{D>C} < E_{BD} < E_D$		

A Comparative Study of Particle Filters and its Variants in Lithium-ion Battery SOH Estimation

Dawei Pan*
College of Information and Communication Engineering
Harbin Engineering University
Harbin, China
pandawei@hrbeu.edu.cn

Hengfeng Li
College of Information and Communication Engineering
Harbin Engineering University
Harbin, China
lihengfeng_77@163.com

Yuchen Song
School of Electronics and Information Engineering
Harbin Institute of Technology
Harbin, China
songyuchen@hit.edu.cn

Abstract—**With the wide application of lithium-ion battery in various fields, the State of Health (SOH) estimation has become a research hotspot for advanced battery management system (BMS). Accurate SOH estimation is helpful to ensure the safe operation of equipment or system in practical applications. Among various lithium-ion battery health diagnosis methods, particle filter and its variants are the mainstream with the significant advantages in non-linear and non-Gaussian system modeling. But in practical applications, the BMS always suffers from the limited power supplication and finite computing resources. Therefore, this paper implemented a comparative study on particle filter (PF) and typical variants, including extended Kalman particle filter (EPF), unscented particle filter (UPF), regularized particle filter (RPF). Through the NASA's battery degradation model, the performance of the above particle filter algorithms is compared and analyzed. The experimental results show that UPF has the highest estimation accuracy, and it is more suitable for the situation with higher prediction accuracy requirements. PF has the least time consumption and is more suitable for on-line health assessment.**

Keywords—lithium-ion battery, state of health, particle filter, extended Kalman particle filter, unscented particle filter, regularized particle filter

I. INTRODUCTION

Lithium-ion battery has been widely used in various industrial applications, with the advantages of high energy density, long cycle life etc. [1]. However, the continuous Li loss during charging and discharging and other irreversible electrochemical reations will gradually degrade the battery performance, which will affect the operational reliability and safety of the host system. Therefore, the health monitoring of lithium-ion battery is one of the key issues to ensure the reliable operation of the system [2]

Battery capacity can be used as a parameter to represent the health status of lithium-ion battery, and the degradation of battery performance can be quantitatively evaluated by capacity modeling [3]. Model-based method have been widely researched for lithium-ion battery health state diagnosis and prognosis [4, 5]. Statistical filtering approach can be applied to model-based methods, such as extended Kalman filter (EKF) [6, 7], unscented Kalman filter (UKF) [8, 9], particle filter (PF) [10], and so on. EKF and UKF can be applied to weakly nonlinear system, but they still cannot meet the requirements of lithium-ion batteries [11]. This is because the lithium-ion battery is a non-linear and non-Gaussian system, which is affected by temperature, various electrochemical reactions and other factors [1]. Adaptive EKF [12] and adaptive UKF [13, 14] have been developed to adapt to the battery system, but their performance may also be influenced

by the characteristics of the lithium-ion battery system. Compared with Kalman Filter and its variants, particle filter can be well applied to such systems without any restrictions on system noise, it approximates the stochastic Bayesian estimation of nonlinear systems by predicting and updating the sampled particle set from the probability density function of the system. Since particle filter was proposed, a variety of improved particle filtering algorithms have been proposed in succession. Chang et al. [15] Optimizes particle distribution to avoid particle degradation. It is also possible to avoid particle degradation by improving resampling. Zhong et al. [16] realizes the state estimation of lithium-ion battery based on unscented particle filter (UPF).

These PF-based lithium-ion battery SOH estimation methods can realize high estimation accuracy. But another issues that should be taken into consideration is that the complexity of PF algorithm is much higher than KF and its variants. For an embedded battery management system (BMS), its power consumption is limited by the host system. At the same time, the computing ability for a BMS with embedded processor is much lower than the computer. Therefore, finding a proper SOH estimation method that balances the estimation accuracy and computing complexity is an urgent issue in practical applications. In this study, we take PF, extended Kalman particle filter (EPF), UPF, regularized particle filter (RPF) as the research object, and compare the accuracy and running time of the four kinds of filters in SOH.

Focusing on the requirements on accuracy and calculating speed of a battery management system(BMS), this work compares the PF and its variants for battery SOH estimation.

(1) In order to evaluate the accuracy of the four different algorithms, the estimation errors of SOH with different particle numbers are compared.

(2) The running time of different particle numbers is measured, and the real-time performance of these four filter algorithms are analyzed.

The purpose of our study is to provide the reference for the selection of filters of SOH estimation in different scenarios. The rest of this paper is organized as follows. The basic principles of standard particle filter is described in Sect. 2. The improved particle filter algorithms are given in Sect. 3. Section 4 mainly shows the experimental results and discusses them. Section 5 concludes the whole paper.

II. GENERAL PROCESS OF STATE OF HEALTH ESTIMATION BASED ON STASTICAL FILTER APPROACH

SOH represents the health status of a battery, which is the ratio of the capacity released by the battery from its full state

---

to its nominal capacity at a certain rate of discharge to the cut-off voltage. It is usually a prediction of the battery life. And with the increase of the number of battery charges and discharges, the capacity of the battery decreases and the internal resistance increases. Therefore, SOH is commonly defined by capacity and internal resistance. The commonly used SOH estimation methods can be divided into three categories: experimental estimation [17], adaptive filtering, and data-driven [18].

The experimental estimation method is a SOH estimation method that requires a large number of experiments in the laboratory and analysis of battery aging behavior. However, it is not applicable because of the high requirements of SOH estimation in laboratory. The data-driven method does not need to understand the working principle and battery model of the battery, only related to the aging data collected. However, the data-driven method is highly dependent on data, and it is difficult to collect enough data in practical applications. The adaptive filtering method can identify the parameters of the battery health state in the model, and then estimate the SOH.PF is its typical algorithm. Since the battery SOH estimation involves solving non-Gaussian problems based on nonlinear time-varying systems. To solve such problems, PF has certain advantages, and the accuracy can approximate the optimal estimation. When a new value is input, the current state is only related to the state of the previous moment, updating the weight and position of particles based on measurements. This study uses capacity to represent the battery performance degradation. Fig. 1 gives the diagram of general flow of SOH estimation based on PF algorithm. The proposed method has the following steps:
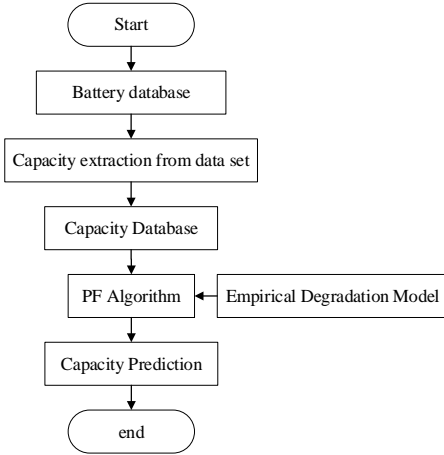
Fig. 1. The schematic diagram of SOH estimation.

Step 1: The battery capacity data is extracted from the battery data and preprocessed.

Step 2: The sample data and prediction objects are selected, and the battery capacity degradation model is established based on the sample data.

Step 3: Initialize the particle filter algorithm and set some related parameters.

Step 4: Particle filter algorithm is used to estimate battery SOH, including particle set initialization, importance sampling, weight calculation, resampling and battery capacity state estimation.

III. PARTICLE FILTER AND ITS VARIANTS

A. Particle Filter

PF is a statistical filtering method of recursive Bayes estimation, which is based on Monte Carlo method to calculate the integral in Bayes estimation. The elemental idea is: by randomly selecting a set of particles with weights in the state space, according to the measurement results, the particles weight is adjusted continuously and state is updated recursively. PF is summarized as follows.

1) *Initialization: by prior probability $p(x_0)$ generate particle set $x_0^i : p(x_0)$.*

2) *Importance sampling: sequential importance sampling (SIS) through particle recommendation distribution $x_k^i : q(x_k^i \mid x_{0:k-1}^i, y_{0:k})$.*

3) *Weight calculation and normalization:*

$$w_k^j \propto \frac{p(y_k \mid x_k^i) p(x_k^i \mid x_{k-1}^i)}{q(x_k^i \mid x_{k-1}^i, y_{1:k})} \tag{1}$$

$$w_k^i = w_k^i \Big/ \sum_{j=1}^{N} w_k^j \tag{2}$$

4) *Resampling:*

$$x_k^i = x_k^j, \sum_{j=1}^{N} w_k^j \geq r_k^i \tag{3}$$

$$\overline{w}_k^i = 1/N \tag{4}$$

5) *State estimation:*

$$\overline{x}_k = \sum_{j=1}^{N} \overline{w}_k^i x_k^i \tag{5}$$

Fig. 2 gives the diagram of flow of standard PF algorithm.
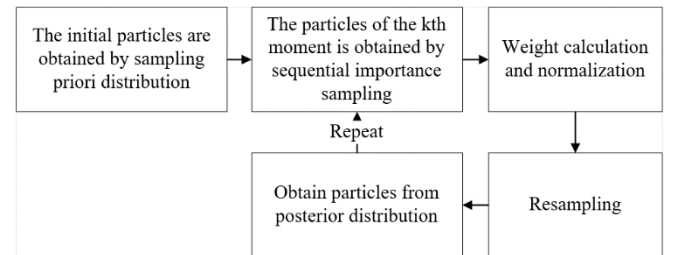
Fig. 2. Standard PF algorithm flow.

To a great extent, the results of state estimation by particle filter are affected by the selection of importance density function and importance sampling. Because prior probability density is chosen as the importance density function by standard PF algorithm, if the system requires high prediction accuracy, it will not be able to obtain ideal results. And the resampling process will continuously delete the weight smaller particles and copy the larger ones. With the repeated recursion, most left particles are those with large weights, which loses the diversity of particles and lead to degradation. In order to solve the above two problems effectively,

improved particle filter algorithms are proposed. Such as improved importance density function: EPF, UPF; improved particle filter algorithm based on resampling technology: RPF.

B. Extended Kalman Particle Filter

As a common filtering method, EKF is suitable for nonlinear systems. The basic principle is to linearize the nonlinear system by first-order Taylor approximation of the nonlinear system. In particle filter algorithm, the importance density function can be generated by EKF. The elemental idea is that the sampled particles are updated by EKF algorithm, and new particles are generated through the importance density function. After the weight is updated, resampling the particle set. The main steps of EPF are summarized as follows:

1) *Update particles with EKF*:

$$x_{k|k-1}^i = f(x_{k|k-1}^i) \tag{6}$$

$$P_{k|k-1}^i = F_k^i P_{k-1}(F_k^i)^T + F_k^i Q_k (F_k^i)^T \tag{7}$$

$$K_k = P_{k|k-1}^i H_k^i [U_k^i R_k (U_k^i)^T + H_k^i P_{k|k-1}^i (H_k^i)^T]^{-1} \tag{8}$$

$$\hat{x}_k^i = \hat{x}_{k|k-1}^i + K_k(z_k - h(\hat{x}_{k|k-1}^i)) \tag{9}$$

$$\hat{P}_k^i = P_{k|k-1}^i - K_k H_k^i P_{k|k-1}^i \tag{10}$$

2) *Update generated particles*:

$$x_k^i : \ q(\hat{x}_k^i \mid x_{k-1}^i, z_k) = N(\hat{x}_k^i \mid \hat{P}_k^i) \tag{11}$$

Fig. 3 gives the diagram of flow of EPF algorithm.
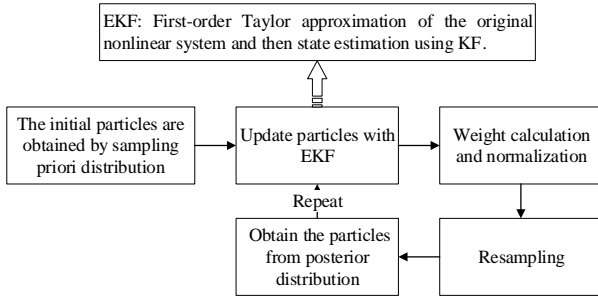
Fig. 3. EPF algorithm flow.

C. Unscented Particle Filter

The principle of UKF is to make the nonlinear equation adaptable for Kalman filter by unscented transformation. Due to the unscented transform is not linearized and the higher-order terms are not ignored, the mean and covariance with high accuracy can be obtained. Therefore, unscented Kalman filter can be used to generate importance density function, update particles, and apply to particle filter. The elemental idea is that the sampled particles are calculated by UKF algorithm, and the mean and variance obtained are used for the next update. The main steps of UKF algorithm to update particles are as follows.

1) *Select particles*:

$$x_{k-1}^{ia} = \left[\overline{x}_{k-1}^{ia} \quad \overline{x}_{k-1}^{ia} \pm \sqrt{n_a + \lambda} \right] \tag{12}$$

Where $n_a = n_x + n_w + n_v$ and $x_{k-1}^{ia} = \left[x_{k-1}^{ix} \quad x_{k-1}^{iw} \quad x_{k-1}^{iv} \right]^T$ is the extended state vector of x_{k-1}^{ia}.

2) *Time update*:

$$x_{k|k-1}^{ix} = f(x_{k-1}^{ix}, u_{k-1}, x_{k-1}^{iw}) \tag{13}$$

$$\overline{x}_{k|k-1}^i = \sum_{j=0}^{2n_a} W_j^{(m)} x_{j,k|k-1}^{ix} \tag{14}$$

$$P_{k|k-1}^i = \sum_{j=0}^{2n_a} W_j^{(c)} \left[x_{j,k|k-1}^{ix} - \overline{x}_{k|k-1}^i \right] \left[x_{j,k|k-1}^{ix} - \overline{x}_{k|k-1}^i \right]^T \tag{15}$$

$$z_{k|k-1}^i = h(x_{k-1}^{ix}, x_{k-1}^{iv}) \tag{16}$$

$$\overline{z}_{k|k-1}^i = \sum_{j=0}^{2n_a} W_j^{(m)} z_{j,k|k-1}^i \tag{17}$$

3) *Measurement update*:

$$P_{z_{k|k-1}z_{k|k-1}} = \sum_{j=0}^{2n_a} W_j^{(c)} \left[\overline{z}_{j,k|k-1}^i - \overline{z}_{k|k-1}^i \right] \left[\overline{z}_{j,k|k-1}^i - \overline{z}_{k|k-1}^i \right]^T \tag{18}$$

$$P_{x_{k-1}z_{k|k-1}} = \sum_{j=0}^{2n_a} W_j^{(c)} \left[x_{j,k|k-1}^{ix} - \overline{x}_{k|k-1}^i \right] \left[\overline{z}_{j,k|k-1}^i - \overline{z}_{k|k-1}^i \right]^T \tag{19}$$

$$K_k = P_{x_{k|k-1}z_{k|k-1}} P_{z_{k|k-1}z_{k|k-1}}^{-1} \tag{20}$$

$$\overline{x}_k^i = \overline{x}_{k|k-1}^i + K_k(z_k - \overline{z}_{k|k-1}^i) \tag{21}$$

$$\hat{P}_k^i = P_{k|k-1}^i + K_k P_{z_{k|k-1}z_{k|k-1}} K_k^T \tag{22}$$

4) *Update Generated Particles*:

$$\hat{x}_k^i : \ q(\overline{x}_k^i \mid \overline{x}_{0:k-1}^i, z_{1:k}) = N(\overline{x}_k^i, \hat{P}_k^i) \tag{23}$$

Fig. 4 gives the diagram of flow of UPF algorithm.
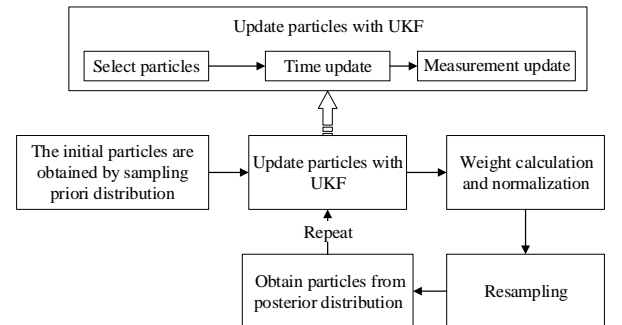
Fig. 4. UPF algorithm flow.

D. Regularized Particle Filter

RPF algorithm is improved based on resampling technique to alleviate the degradation of particle diversity. The sample of the algorithm is extracted from

$$p(x_k \mid z_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i K_h(x_k - x_k^i), \qquad (24)$$

$K_h(x)$ is the kernel density function and satisfies

$$K_h(x) = \frac{1}{h^{n_x}} K(\frac{x}{h}), \qquad (25)$$

where n_x is the dimension of state space x; $h > 0$ is the core bandwidth, $K(\cdot)$ is a standard kernel density function and satisfies

$$\int x K(x)dx = 0, \int \|x\|^2 K(x)dx < \infty. \qquad (26)$$

when the particle is equal weight, the optimal kernel density function is Epanechnikov kernel,

$$K_{opt} = \begin{cases} \dfrac{n_x + 2}{2c_{n_x}}(1 - \|x\|^2), \|x\| < 1 \\ 0 \qquad\qquad , \text{ other} \end{cases}. \qquad (27)$$

where c_{n_x} is the volume of the unit sphere on R^{n_x} and n_x is the dimension of the distribution. The optimal bandwidth choice is

$$h_{opt} = AN_s^{1/n_x + 4}, \qquad (28)$$

in the formula,

$$A = [8c_{n_x}^{-1}(n_x + 4)(2\sqrt{\pi})^{n_x}]^{1/(n_x + 4)}, \qquad (29)$$

Fig. 5 gives the diagram of flow of RPF algorithm.
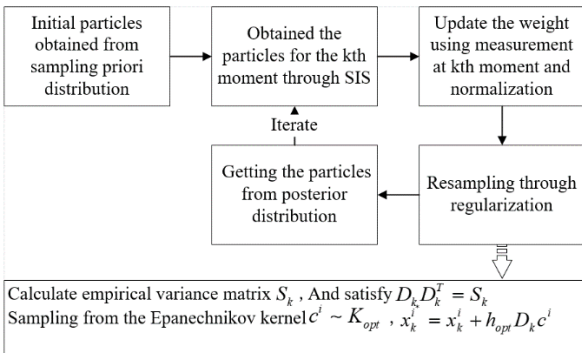
Fig. 5. RPF algorithm flow.

E. Chapter Summary

According to the structure of the particle filter algorithm, we can divide it into five steps: initialization, importance, weight calculation and normalization, resampling and state

estimation. Its variants are improved in different steps. UPF and EPF are improved in the second step, and EKF and UKF are used to generate the importance density function for sampling. RPF has improved in the fourth step, using regularization to resample, alleviating the problem of particle diversity degradation. The following table is a comparison of algorithm steps.

TABLE I.   COMPARISON OF ALGORITHM STEPS

Filter	Algorithm Step				
	Step 1	Step 2	Step 3	Step 4	Step 5
PF		SIS			
EPF	The same	EKF	The same	The same	The same
UPF		UKF			
RPF		SIS		Regular resampling	

IV. VARIANTION AND COMPARISON

The experimental data used in this paper are from the Center for Advanced Life Cycle Engineering (CALCE) of the University of Maryland. The rated capacity of the lithium-ion battery used is 1.1Ah, and the failure threshold of the battery is 0.88Ah, which is the battery capacity when SOH = 80%.

In this experiment, the performance of the algorithm is investigated by the degradation model of battery capacity. Generally, the following equation can be used to describe the degradation process of lithium-ion battery capacity:

$$C_{k+1} = \eta_c C_k + \beta_1 \exp(-\beta_2/\Delta t_k) \qquad (30)$$

where C_k shows the charging capacity of the k_{th} cycle, represents the interval between kth cycle and the $k+1_{th}$ cycle, β_1 and β_2 are undetermined parameters. The state transition equation and observation equation are established according to the model. In the formula, $\beta_1 = -0.3$, $\beta_2 = 5$.

state transition equation:

$$C_{k+1} = 0.997 C_k - 0.3\exp(-5/\Delta t_k) + \omega_k \qquad (31)$$

observation equation:

$$y_k = C_k + \upsilon_k \qquad (32)$$

In the experiment, the number of simulation steps is the whole battery capacity data (T=132), and the value of particles N are 50, 100, 200, 300, 400, and 500, respectively. Each value was run 15 times to avoid the influence of randomness on the results. The average of 15 results was the final result. PF, EPF, UPF and RPF are used for state estimation and explore the performance of four kinds of particle filter.

This work introduced two evaluation criteria: to quantitatively evaluate the performance of different PF algorithms, namely, the Root Mean Square Error (RMSE) and running time. RMSE is used as the standard of filter estimation performance, and the results are shown in Table 2. and Fig. 6. Take the running time (t) as the standard of real-time

performance, the results are shown in Table 3. and Fig. 7. The definition of these two parameters are as follows，

$$RMSE = \sqrt{\frac{1}{T}\sum_{k=1}^{T}(\hat{x}_k - x_k)^2} \qquad (33)$$

$$t = \sum_{k=1}^{T} t_k \qquad (34)$$

TABLE II. RMSE OF FOUR KINDS OF FILTERS UNDER DIFFERENT CONDITIONS

Particle Numbers	RMSE			
	PF	EPF	UPF	RPF
N=50	0.0203	0.019	0.0186	0.0196
N=100	0.0161	0.0155	0.0154	0.0155
N=200	0.0104	0.0102	0.0101	0.0103
N=300	0.0086	0.0085	0.0083	0.0086
N=400	0.0075	0.0074	0.0073	0.0074
N=500	0.0067	0.0066	0.0064	0.0066
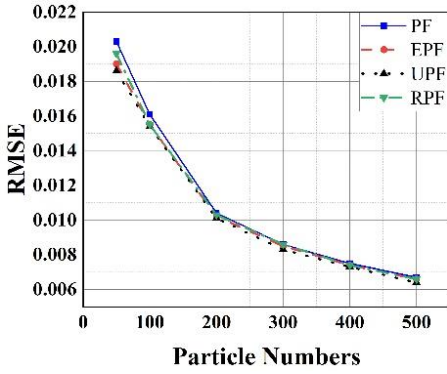
Fig. 6. RMSE of four kinds filters under different conditions.

TABLE III. RUNNING TIME OF FOUR KINDS OF FILTERS UNDER DIFFERENT CONDITIONS

Particle Numbers	Running Time(t)			
	PF	EPF	UPF	RPF
N=50	0.1572	0.2109	0.4359	3.9788
N=100	0.2597	0.3955	0.8681	15.5763
N=200	0.5533	0.7927	1.6889	62.7772
N=300	0.768	1.1568	2.559	139.9288
N=400	1.0921	1.5925	3.3943	259.1386
N=500	1.2039	1.8642	4.1376	374.762
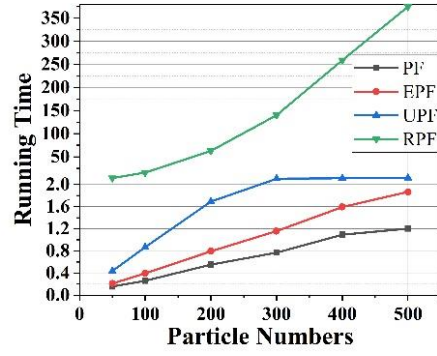
Fig. 7. Running time of four kinds filters under different conditions.

As shown in Table 2. and Fig. 6, UPF has a better performance under the same particle number, and its RMSE is smaller than other particle filter algorithms in most cases. It shows that when UPF is used in battery degradation system, its estimation accuracy is high, and it is more suitable for battery SOH estimation. PF and RPF are not good as UPF, but as shown in Table 3. and Fig. 7, the state estimation of PF has the shortest running time and the best real-time performance under the same number of particles. This is because pf occupies less RAM and the algorithm is simple. RPF takes the longest time because of its high computational complexity.

As shown in Table 2. and Fig. 6, in the case of different particle number, the state estimation performance of all particle filtering algorithms is more accurate as the number of particles increases. Take PF as an example, when N=500, the estimated accuracy is three times higher than N=50. However, as shown in Table 3. and Fig. 7, when the number of particles in the PF algorithm increases to 500, the running time is nearly eight times longer than when $N=50$. It shows that the filtering accuracy is obtained at the expense of real-time performance. Comprehensive consideration, when the number of particles $N=200$, the estimation accuracy and running time are better.

Due to space limitations, we cannot list the particle distribution in each case, only the most comprehensive particle distribution is listed, that is, UPF at $N=200$. Fig. 8 gives the diagram of UPF particle set distribution.

(a) N=200,T=10.   (b) N=200,T=50

(c) N=200,T=100.
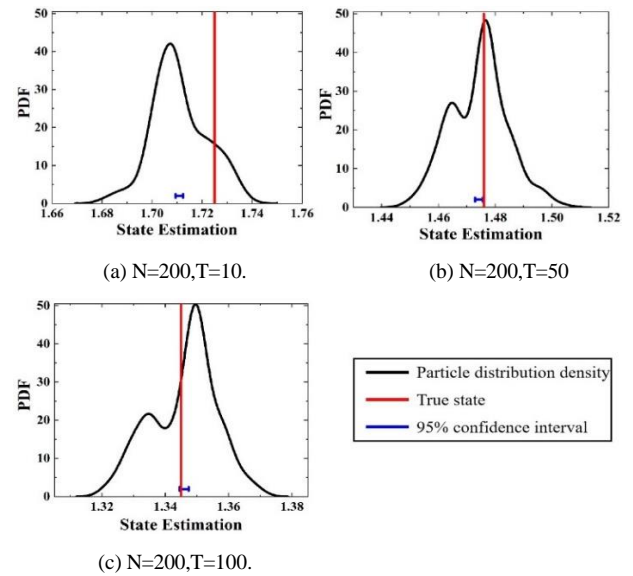
Fig. 8. The schematic diagram of UPF particle set distribution.

As can be seen from the (a) (b) (c) diagram in Fig. 8, with the increasing number of iterations, the range of values of the particle set is shrinking and getting closer to the true state. It is shown that resampling solves the problem of particle degeneration. With iteration, And the resampling process will continuously delete the weight smaller particles and copy the larger ones. Replace $\left\{ x_k^i \right\}_{i=1}^{N}$ with the new sampling value $\left\{ x_k^{i*} \right\}_{i=1}^{N}$, satisfy $p\left\{ x_k^{i*} = x_k^j \right\} = w_k^j$, and update the $w_k^i = N^{-1}$.

Finally, based on the results of this study and the nature of particle filter, the SOH accuracy, running time, algorithm complexity and RAM are compared. Radar diagrams are made as shown in Fig. 9, and four particle filters are compared.
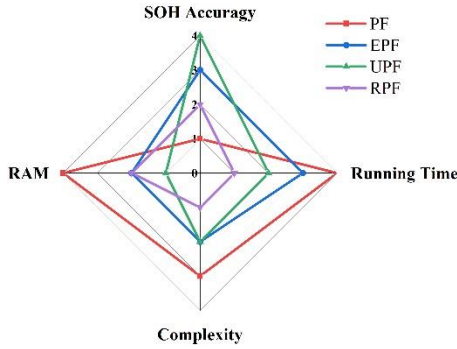
Fig. 9. The schematic diagram of algorithm comparison.

V. CONCLUSION

The above four particle filtering algorithms have realized the SOH estimation of lithium-ion batteries to different degrees. As for PF, the algorithm is relatively simple, although the estimation accuracy is not high enough, the real-time performance is the best. As for EPF, because EKF introduces too many errors in the process of model linearization and Gaussian assumption, taking full advantage of the latest observations, but its improvement is not as good as UPF. The reason why UPF can achieve such a good effect is mainly due to the following two advantages: firstly, it makes full use of the latest observed data; Secondly, the mean and variance of this method can be exact to third-order Taylor expansion, whereas EPF has only first-order. RPF realizes the continuous approximation of the discrete distribution, so that the sampling part can be carried out in the continuous distribution, thus increasing the diversity of particles. Although the accuracy of filter estimation is decreased, it is still a feasible filtering method.

REFERENCES

[1] D. Liu, J. Zhou and L. Guo, "Survey on lithium-ion battery health assessment and cycle life estimation," Yi Qi Yi Biao Xue Bao/Chinese Journal of entific Instrument, vol. 36, no. 1, pp. 1–16, January 2015.

[2] R. Xiong, Y. Zhang, J. Wang, H. He, S. Peng, and M. Pecht, ''Lithium-ion battery health prognosis based on a real battery management system used in electric vehicles,'' IEEE Transactions on Vehicular Technology, vol. 68, no. 5, pp. 4110–4121, May 2019.

[3] J. Zhang and J. Lee, "A review on prognostics and health monitoring of li-ion battery," Journal of Power Sources, vol. 196, no. 15, pp. 6007–6014, 2011.

[4] Y. Song, D. Liu, L. Li and Y. Peng, "Lithium-Ion Battery Pack On-Line Degradation State Prognosis Based on the Empirical Model," International Conference on Sensing, Diagnostics, Prognostics, and Control, Xi'an, China, pp. 402-407, 2018.

[5] H. Meng and Y. F. Li, "A review on prognostics and health management (PHM) methods of lithium-ion batteries," Renewable & Sustainable Energy Reviews, vol. 116, 2019.

[6] C. Jiang, A. Taylor, C. Duan and K. Bai, "Extended Kalman Filter based battery state of charge(SOC) estimation for electric vehicles," IEEE Transportation Electrification Conference and Expo, pp. 1-5, 2013.

[7] J. Huang, Z. Yang, Y. Zhiping, M. Cao and J. Chang, "Parameter identification and SOC estimation of li-ion power battery in hybrid electric vehicle," journal of nanchang university(engineering & technology), pp. 62-69, 2019.

[8] P. Shi and Y. Zhao, "Application of Unscented Kalman Filter in the SOC Estimation of Li-ion Battery for Autonomous Mobile Robot," IEEE International Conference on Information Acquisition, pp. 1279-1283, 2006.

[9] T. Wang, S. Chen, H. Ren and Y. Zhao, "Model-based unscented kalman filter observer design for lithium-ion battery state of charge estimation," International Journal of Energy Research, vol. 42, no. 4, pp. 1603-1614, 2018.

[10] G. Dong, Z. Chen, J. Wei, and Q. Ling, ''Battery health prognosis using brownian motion modeling and particle fifiltering,'' IEEE Transactions on Industrial Electronics, vol. 65, no. 11, pp. 8646–8655, November 2018.

[11] G. Dong, Z. Chen, J. Wei and P. Wang, "An online model-based method for state of energy estimation of lithium-ion batteries using dual filters," Journal of Power Sources, vol. 301, no. 1, pp. 277-286, 2016.

[12] R. Xiong, H. He, F. Sun and K. Zhao, "Evaluation on State of Charge Estimation of Batteries With Adaptive Extended Kalman Filter by Experiment Approach," IEEE Transactions on Vehicular Technology, vol. 62, no. 1, pp. 108-117, January 2013.

[13] L. Huang, Y. Guo and Z. Zhao, "Study on prediction algorithm of AUKF for the lithium-ion battery SOC," IEEE International Conference on Communication Problem-Solving, pp. 608-610, 2015.

[14] E. Hou, X. Qiao and G. Liu, "SOC estimation for power lithium-ion battery based on AUKF," International Conference on Artificial Intelligence and Engineering Applications, 2016.

[15] Y. Chang and H. Fang, ''A hybrid prognostic method for system degradation based on particle fifilter and relevance vector machine,'' Reliability Engineering System Safety, vol. 186, pp. 51–63, June 2019.

[16] L. Zhong, C. Zhang, Y. He and Z. Chen, "A method for the estimation of the battery pack state of charge based on in-pack cells uniformity analysis," Applied Energy, vol. 113, pp. 558-564, 2014.

[17] R. Xiong, L. Li and J. Tian, "Towards a smarter battery management system: A critical review on battery state of health monitoring methods," Journal of Power Sources, vol. 405, pp. 18-29, November 2018.

[18] D. Zhou, X. Song, W. Lu and P. Fu, "Real-time SOH estimation algorithm for lithium-ion batteries based on daily segment charging date," Proceedings of the CSEE, vol. 39, pp. 105-111, January 2019.

The Mathematical Construction of The Battery Mechanism Function

Li Wen
dept.of Thermal Science and Energy
Engineering
University of Science and Technology
of China
Hefei, China
liwen96@mail.ustc.edu.cn

Mao Lei
dept.of Precis Machinery & Precis
Instrumentat
University of Science and Technology
of China
Hefei, China
leimao82@ustc.edu.cn

Zhang Chen
dept.of Precis Machinery & Precis
Instrumentat
University of Science and Technology
of China
Hefei, China
zhang985@mail.ustc.edu.cn

Xiaofang Cheng*
dept.of Thermal Science and Energy
Engineering
University of Science and Technology
of China
Hefei, China
xfcheng@ustc.edu.cn

Abstract—**It is helpful to guide the development and application of batteries to establish a correct volt ampere function mechanism model. But, two scientific issues need to be solved: First, how many mechanisms battery has; Second, how to establish the mechanism model alone, in the case of overlapping of these mechanisms. There are linear and nonlinear states in volt ampere characteristics. The mathematical property of monotonic decreasing of volt-amperes characteristics indicates that battery is of only three kinds of mechanisms. Without changing the basic form of volt ampere function, and under the principle of considering the mechanism function's working region, we propose a mechanism function which meets the monotonic decreasing of the voltammetric curve of battery. The proposed mechanism function is based on the derivative law of each mechanism function in the voltammetric function of battery. By using the voltammetric data, the obtained mechanism function can accurately predict the potential (current or voltage) when the independent variable of battery is zero, and also it lays a theoretical foundation for the internal working mechanism of battery**.

Keywords—volt-ampere characteristics, battery mathematical model, mechanism function

I. INTRODUCTION

Battery, a kind of chemical energy, can realize the conversion between electric energy and chemical energy, and also is one of the effective ways to solve the energy crisis. Therefore, battery is of great significance in the field of power energy and energy storage. The curtain of battery research was opened since fuel cell was invented by Britain's William Grove in 1839, and photovoltaic effect discovered by Alexander-Edmond Becquerel concurrently[1-3].

The volt ampere characteristic functions (VACFs)of different types of batteries has been widely used in the research of batteries, and the research on the nonlinear volt ampere characteristic function of batteries is in full swing. Many scholars use modeling and simulation methods to get the volt ampere function of battery. Cheng[4] et al. established the mathematical model of fuel cell mechanism function, based on least square method and basic electrochemical equation; using MATLAB/Simulink, Jia[5] et al. developed a new dynamic model of proton exchange membrane fuel cell(PEMFC). The simulation result shows the model is effective and operational and that the structure of PEMFC is optimized to improve its performance. The equation developed by Amphlett [6] has accepted its compactness in predicting the behavior of PEMFCs under different operating conditions. However, due to the properties of complex system, multivariable and strong coupling of different types of batteries, it is difficult to model those properties for battery design and performance evaluation[7]. The explicit expression of VACFs is the key of battery theoretical research. At present, the expression of battery's VACFs is various under different working conditions, and the research's conclusion has not reached the effect of guiding practice.

From the mathematical point of view, this paper analyzes and deduces the specific expression of battery's VACFs, and gives its expression which can guide the practice.

II. BASIC PHENOMENA OF BATTERY VOLT-AMPERE CHARACTERISTICS

In physical experiments, the mathematical expression of VACFs can be shown as following

$$y = f(x) \qquad (1)$$

Rectangular coordinate system constructed by current and voltage is called volt ampere coordinate system [10,11]. According to the value of output current and output voltage , the curve of VACFs can be drawn in the volt ampere coordinate system, and the curve is located in the first quadrant of the coordinate system, namely $x \geq 0, f(x) \geq 0$. Experiments show that [8,9] the volt-ampere function has the characteristic of monotone decreasing, which can be expressed mathematically as $f_k'(x) < 0$. The VACFs is the following.

$$f(x) = A - \sum_{k=1}^{N} f_k(x) \geq 0 \left(0 \leq f_k(x) \leq A \right) \qquad (2)$$

Where x ,independent variable, means current or voltage; $f(x)$ is dependent variable, and it means voltage or current; A , potential parameter, provides the electromotive force for the voltage source battery, and the current for the

---

current source battery; $f_k(x)$ is the battery mechanism function. However, those mechanism functions can attenuate power supply parameters ,so the curve of VACFs presents monotonic decreasing $f'(x) < 0$; N is the specific number of mechanism functions.

VACFs should obviously meet the following requirements.

$$f(0) = A - \sum_{k=1}^{N} f_k(0) = A \qquad (3)$$

$f_1(0) + f_2(0) + L + f_N(0) = 0$, whose solution is not closed, but the experimental data of each mechanism function cannot be separated alone. N , and the specific form of $f_k(x)$ cannot be determined by (3). That is the defect in the expression form of explicit function summation of VACFs.

The properties of the curve of VACFs determine N , $f(x): N$.

$$\sum_{k=1}^{N} f_k(x) \Rightarrow f_k(x), k = 1, 2, ...N \qquad (4)$$

By separating the functions from each other in $f_k(x)$, the specific number and form of VACFs can be accurately expressed.

$$f_1(x) + f_2(x) + L + f_N(x), k = 1, 2, ...N$$
$$\rightarrow f_1(x),\ f_2(x), L\ ,\ f_N(x), k = 1, 2, ...N \qquad (5)$$

Therefore, the main task of this paper is to find the separation method of VACFs, determine the specific number and form of mechanism function respectively, and express the VACFs accurately.

III. CONSTRUCTION OF THE CORRECT BATTERY VOLT AMPERE FUNCTION

A. *Establishment of the specific number of battery mechanisms*

The curve of VACFs can be expressed by $f(x)$ in the plane coordinate system, and $f_k'(x) < 0$. The range of X is $0 \le x \le x_{max}$,and the range of $f(x)$ is $0 \le f(x) \le A$.

The mathematical properties of VACFs is $f'(x) < 0$,and the following volt-ampere function can be obtained by definite integral.

$$f'(x) = -B \xrightarrow{\text{definite integral}} f(x) = A - Bx \qquad (6)$$

After nonlinear treatment of (6), the nonlinear function can be expressed as

$$f(x, n) = A - Bx^n \ge 0\,(A, B, n > 0)$$
$$\rightarrow \begin{cases} n = 1, \text{linear} \\ n \ne 1, \text{nonlinear} \end{cases} \qquad (7)$$

The volt-ampere function of (7) is suitable for all types of batteries and called general volt ampere function(GVAF) . n is called weighting factor of nonlinear function, and

parameters (A, B, n) still need to be given clear physical connotation.

Analysis of GVAF begins from the slope of any point and the overall trend in the volt ampere curve. The first derivative corresponds to any point on the curve, and the second derivative reflects the whole trend of the curve.

To take the first derivative of the GVAF, thus

$$f'(x) = -nBx^{n-1} < 0 \qquad (8)$$

The first derivative of the GVAF meets the requirements of $f'(x) < 0$.

To take the second derivative of the GVAF, thus

$$f''(x) = -n(n-1)Bx^{n-2}$$
$$\Rightarrow \begin{cases} \xrightarrow{n=1} f''(x) = 0 \\ \xrightarrow{n \ne 1} \begin{cases} \xrightarrow{n>1} f''(x) < 0 \\ \xrightarrow{n<1} f''(x) > 0 \end{cases} \end{cases} \qquad (9)$$

In (9), the symbolic state of exponent n determines the opening direction of GVAF, and the opening direction reflects the battery mechanism function. Therefore, the state of n corresponds to the form of the battery mechanism function, and that can be used as a criterion for the mechanism function.

Table 1 shows the battery mechanism function only has three different forms. The first mechanism corresponds to zero, the second corresponds to less than zero, and the last corresponds to greater than zero. As shown in Table 1.

TABLE 1 RELATIONSHIP BETWEEN WEIGHTED PARAMETERS n AND CELL MECHANISM

Explicit expression of volt ampere function	Numerical state of independent variable index	Second derivative case	Mechanism model of battery
$f(x, n) = A - Bx^n$	$n = 1$	$f''(x) = 0$	First mechanism model
	$n > 1$	$f''(x) < 0$	Second mechanism model
	$n < 1$	$f''(x) > 0$	Third mechanism model

Table 1 reveals three kinds of battery mechanism functions. Further analysis, however, is expected whether those functions can be used as battery mechanism functions.

By decomposing (7), it can be achieved.

$$f(x) = A - Bx - Bx(x^{n-1} - 1)$$
$$= A - f_1(x) - f_2(x) \ge 0,\ x \ge 0$$
$$\Rightarrow \begin{cases} n = 1 \rightarrow f_1(x) = Bx \ge 0, x \ge 0 \\ n \ne 1 \rightarrow \begin{cases} n > 1 \xrightarrow{f_2''(x)<0} f_2(x) = Bx(x^{n-1} - 1) \ge 0, x \ge 1 \\ n < 1 \xrightarrow{f_3'(x)>0} f_3(x) = Bx(x^{n-1} - 1) \ge 0, x \le 1 \end{cases} \end{cases}$$
$$\qquad (10)$$

Three kinds of battery mechanism functions can be deduced by integration of (10). The range of x is different, and the functions are of the subsection and exist independently. However, those three different battery mechanism functions actually only have two different forms: Bx and $Bx(x^{n-1} - 1)$ .When $n < 1$, $x \le 1$, the independent

TABLE2 DERIVATIVE CHARACTERISTICS OF THE BATTERY MECHANISM FUNCTION

Battery mechanism	First mechanism model	Second mechanism model	Third mechanism model
First derivative	$f_1'(x) > 0$	$f_2'(x) > 0$	$f_3'(x) > 0$
Second derivative	$f_1''(x) = 0$	$f_2''(x) < 0$	$f_3''(x) > 0$
Mechanism function form	$f_1(x) = Bx$	$f_2(x) = \exp(x)$	$f_3(x) = \ln(x)$
Mechanism function requirement	$f_1(x) = Bx$	$f_2(x) = \exp(x) - 1$	$f_3(x) = \ln(1 + x)$

$f_3(x) = Bx(x^{n-1} - 1)$, does not meet $0 \leq x < x_{\max}$.

Apparently, GAVF obtained by integration cannot be used to construct the concrete form of mechanism function.

B. *Construction of the specific mechanism function form of battery $f_k(x)$*

From the above, the battery mechanism function is $\left[f_k(x), k = 1, 2, 3 \right]$. Before the specific form construction of $f_k(x)$, it should be separated firstly.

Mathematical deduction is carried out from the derivative of VACF. Then, the proper function form can be gotten according to the derivative law of mechanism function. The original function of VACF $f_k(x)$, the first and second derivative of VACF $\left[f_k'(x), f_k''(x) \right]$, all can be shown respectively. Thus, the separation of $f_k(x)$ is completed.

$$f'(x) = -\sum_{k=1}^{n} f_k'(x) < 0 \rightarrow f_k'(x) > 0$$

$$f''(x) = -\sum_{k=1}^{n} f_k''(x) \begin{cases} > 0, 0 \leq x < x_1 \leftarrow n < 1 \\ = 0, x_1 \leq x \leq x_2 \leftarrow n = 1 \\ < 0, x_2 < x \leq x_{\max} \leftarrow n > 1 \end{cases} \quad (11)$$

Comparing the results of (11) with the conditions $f_k(x) \geq 0, x \geq 0$, it shows only $n = 1$, $f_1(x) = Bx \geq 0, x \geq 0$ can meet the conditions.

$$f_1(x) = Bx \geq 0, B > 0, x \geq 0 \quad (12)$$

Formula (12) meets $f_1'(x) > 0, f_1''(x) = 0, n = 1$.

Besides, the non-linear mechanism function $(n \neq 1)$ needs to be reconstructed. When $n \neq 1$, there are two kinds of mechanism function: $f_2(x)$ and $f_3(x)$.

In the light of $f''(x) > 0$ and $f''(x) < 0$, the inverse proportional function, trigonometric function, logarithmic function, and exponential function in the nonlinear function need to meet the conditions that the first derivative is less than zero and the second derivative is greater than zero and less than zero. But, the curve of inverse proportional function $f(x) = \dfrac{k}{x}$ has no intersection with the coordinate axis, and that function does not meet the battery

characteristics. The trigonometric function is a periodic function, which does not conform to the monotonicity of battery mechanism function and the unevenness of the volt-ampere curve. On the contrary, both logarithmic function and exponential function meet the conditions. Meanwhile, previous studies have also confirmed that the battery mechanism function does contain the logarithmic and exponential functions[13-22].

While constructing a mathematical function, the form of function should not be adjusted to meet the experimental data of that type, because functions are objective reflections of natural laws in mathematics. Only in this way, can it be ensured that the physical meaning of the functional form will not be distorted by the change of the functional form. Also, the selected mathematical functions should be matched according to the dimension and physical meaning.

The mechanism function of $f_2(x)$ should meet $f_2'(x) > 0, f_2''(x) > 0, n > 1$. Choosing exponential function with natural constant e as base, $f_2(x) = \exp(x)$.

$$f_2'(x) = \exp(x) > 0, f_2''(x) = \exp(x) > 0 \quad (13)$$

To meet $f_k(0) = 0$, the mechanism function is expressed as $f_2(x) = \exp(x) - 1$. After configuring dimensions, the function is the following.

$$f_2(x) = C_0 \left[\exp(C_1 x) - 1 \right] \geq 0, x \geq 0 \quad (14)$$

The mechanism function of $f_3(x)$ should meet $f_3'(x) > 0, f_3''(x) < 0, n < 1$. Choosing a logarithmic function with the natural constant e as the base, $f_3(x) = \ln(x)$.

$$f_3'(x) = \frac{1}{x} > 0, f_3''(x) = -\frac{1}{x^2} < 0 \quad (15)$$

To meet $f_k(0) = 0$, the mechanism function is expressed as $f_3(x) = \ln(1 + x)$. After configuring dimensions, the function is the following.

$$f_3(x) = D_0 \ln(1 + D_1 x) \geq 0, x \geq 0 \quad (16)$$

The volt-ampere characteristics of battery mechanism function should meet $f'(x) < 0$ and $\sum_{k=1}^{n} f_k(0) = 0$, but the derivative properties are quite different

$f_1''(x) = 0, f_2''(x) > 0, f_3''(x) < 0$. The derivative properties of battery mechanism function can be sorted out as shown in Table 2.

The battery mechanism function which meets the conditions has been gotten.

$$f(x) = A - \sum_{k=1}^{3} f_k(x) \qquad (17)$$
$$= A - Bx - C_0\left[\exp(C_1 x) - 1\right] - D_0 \ln(1 + D_1 x)$$

So far, the battery mechanism function has been constructed and is applicable to all types of batteries.

IV. APPLICATION AND VALUE OF THE UNIVERSAL VOLTAMMETRY FUNCTION

There exists a theoretical potential value: $A_{\text{theoritical}}$ in the cell (for example, chemical electromotive force of chemical cell, photoelectric conversion efficiency of photovoltaic cell), which usually cannot reach the theoretical potential value; so, the applied potential value $A_{\text{experimental}}$ of battery can be adopted in the experiment. The difference between the two potential values $A = A_{\text{theoritical}} - A_{\text{experimental}}$ can be used to effectively evaluate the state of the battery electrode.

However, the battery data $\left(A_{\text{experimental}}, x_{\max}\right)$ on the coordinate axis in the voltammetric coordinate system is always difficult to measure, so it is impossible to measure $A_{\text{experimental}}$ effectively.

A. Determination and value of the parameters of the universal voltammetry function

There exist six parameters in GVAF in (17). Among those parameters, A is the power parameter, and is the key value in the battery. Those parameters are solved in a closed form $\left(A, B, C_0, C_1, D_0, D_1\right)$. The method of determining power parameter A is the following, when $x = 0$ and $x = x_{\max}$.

$$\begin{cases} f(0) = A_{\text{experimental}} - \sum_{k=1}^{3} f_k(0) = A \\ f(x_{\max}) = A_{\text{experimental}} - \sum_{k=1}^{3} f_k(x_{\max}) = 0 \end{cases} \qquad (18)$$

In the measurement of the battery volt-ampere characteristics, it is extremely difficult to conduct the zero value phenomenon and achieve the result of $V = E$ when the current of the voltage source battery is 0. As for the current source battery, it is also extremely difficult to achieve the result of $I = I_s$ in the experimental measurement, when the voltage is 0. Therefore, it is always a difficult problem to measure the exact value of A of open circuit voltage or short circuit current in battery experiments.

From the theoretical level of battery, $f(0) = A_{\text{theoritical value}}$, and from the aspect of experiment,

$f(0) = A_{\text{experimental value}}$. When the current is zero, $A_{\text{theoritical value}}$ and $A_{\text{experimental value}}$ are not equal.

Then, the theoretical research of battery mechanism function can judge the difference between the actual value and the theoretical value of battery open-circuit voltage. According to $\Delta A = A_{\text{theoritical value}} - A_{\text{experimental value}}$, the error between those two different values is determined.

B. Application of nonlinear batteries

Fuel cell is the most typical non-linear cell, and its curve of volt-ampere characteristic is the most complex curve at present that can reveal three working mechanisms inside the cell: activation polarization loss, ohm polarization loss and concentration polarization loss[23,25,27-29], as shown in Fig.1. Taking fuel cell as an example. From the perspectives of the first and second derivatives and the voltage value when the current is 0, this paper conducts a comparative study between the model of GVAF and the various V-I function models used widely in the analysis and research of proton exchange membrane fuel cells. The chemical composition of fuel cell is expressed in the form of quantity or mathematical function, so it is unnecessary to distinguish the chemical composition of fuel cell.

In[30], an improved mass transport item model is proposed for theoretical analysis of fuel cells.

$$U = E - A\ln(\frac{I + I_0}{I_0}) - m\exp(nI) - IR \qquad (19)$$

In[31], the model is constructed after considering the leakage of current density.

$$U = E - X\ln(\frac{I + I_{in}}{I_0}) + Y\ln(1 - \frac{I + I_{in}}{I_0}) - (I + I_{in})R \qquad (20)$$

In[32], the model of activation loss and mass transport loss expressed by logarithmic function and exponential function.

$$U = E - b\lg\frac{I}{I_L} - m\exp(nI) - RI \qquad (21)$$

In[33], a model considering anode and cathode is proposed.

$$U = E - \left[b_a\ln(\frac{I}{I_{o.a}}) + b_c\ln(\frac{I}{I_{o.c}})\right] - B\ln(\frac{I_L}{I_L - I}) - IR \qquad (22)$$

In[34], the model of activation loss and mass transport loss expressed by logarithmic function.

$$U = E - \frac{RT}{F}\ln\left[\frac{I}{2I_0} + \sqrt{1 + (\frac{I}{2I_0})^2}\right] - \frac{RT}{nF}\ln(1 - \frac{I}{I_L}) - IR \qquad (23)$$

From the first and second derivatives of the volt ampere function and the voltage value when the current is 0, five different functions can be obtained from the above five common fuel cell models. The first derivatives of (22) and (23) are greater than, equal to and less than zero. They do not conform to the characteristic of monotonic decreasing of battery volt-ampere function. When the current is 0, the voltage values in the formulas from (19) to (23) are not the open circuit voltage E, or even infinite voltage value is produced. All does not conform to the actual experimental measurement. Second derivative reflects the opening

direction of battery volt-ampere characteristic curve, and the opening direction of fuel cell volt-ampere characteristic curve is not a single opening direction mechanism. And the second derivative of volt ampere function of fuel cell in (19) and (23) is only less than 0. So, the theory cannot fully reflect the battery's internal mechanism.

Via the analysis of the common functions from the first and second derivative and the voltage value when the current is 0, it can be concluded that the existing theoretical function form of fuel cell is the result of fitting the experimental data on the basis of experiment.

Therefore, there is no sufficient basis to add or reduce the constant term, so that the voltage value of the fuel cell volt ampere function does not appear infinite when the current is 0.It is not in line with the actual situation, and makes the first derivative of the volt ampere function meet the characteristics of monotonic decline.
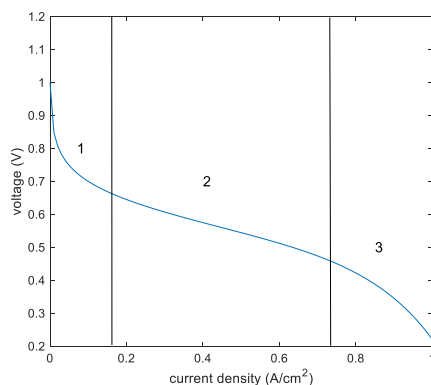
Fig1.Volt-ampere Characteristic Curve of Fuel Cell

When there is a big difference in the fitting of fuel cell volt ampere function to experimental data, it will try to change the symbol in front of the independent variable for a new data fitting. In order to meet the experimental data, the symbol in front of the independent variable is changed. Thus it changes the unequal sign direction of the second derivative of the fuel cell mechanism function. In essence, the physical connotation of battery mechanism function is changed.

V. CONCLUSION

This paper presents a new type of battery volt-ampere function model $f(x) = A - Bx^n$. From the mathematical level, the number of battery internal mechanism is determined and the correct battery mechanism function is established as

$$f(x) = A - Bx - C_0 \left[\exp(C_1 x) - 1 \right] - D_0 \ln(1 + D_1 x).$$

In the process of construction, monotonic decreasing of battery volt ampere curve ,the most basic characteristic, has been used.

By establishing the battery mechanism function in physical layer, it is possible to apply the general volt ampere characteristics of battery to the theoretical level.

The value range of weighting factor n determines the matching of specific mechanism functions. The division of three mechanism functions by the weighting factor n corresponds to the second derivative of the explicit function of the general volt ampere characteristics. For all types of batteries, there is a common feature, that is, volt ampere curve

shows monotonic decreasing and the opening of power curve is downward. Consequently, there lives general volt ampere characteristic function whose complete expression can be obtained by combining mathematical deduction with physical properties .The general function is applicable to all types of batteries.

So, the open circuit voltage and short circuit current can be calculated accurately with the aid of the volt ampere data of battery. The range of exponent x of battery mechanism function corresponds to the three working mechanisms of fuel cell. The specific value of open-circuit voltage can be obtained through the general volt ampere characteristic function . Thereby, the establishment of general volt ampere function lays a foundation and also provides great convenience for the theoretical research of battery.

REFERENCES

[1] Liu Dehui, Liu Xirong. Circuit Basis (in Chinese)[M]. China Water Resources and Hydropower Press, 2004

[2] James Lamir, Andrew Dicks. Fuel Cell System-Principle Design Application [M]. Beijing. Science Press, 2006.

[3] Wang Linshan, LI Ying. Fuel Cell (second edition) (in Chinese)[M]. Beijing: Metallurgical Industry Press,2005.

[4] Xue XD, Cheng KWE, Sutanto D. Unified mathematical modelling of steady state and dynamic voltage–current characteristics for PEM fuel cells. Electrochim Acta 2006;52(3):1135–44

[5] Jia J, Wang Y, Li Q, Cham Y, Han M. Modeling and dynamic characteristic simulation of a proton exchange membrane fuel cell. IEEE Trans Energy Convers 2009;24(1):283–91.

[6] Amphlett JC, Baumert RM, Mann RF, Peppley BA, Roberge PR, Harris TJ. Performance modelling of the Ballard Mark IV solid polymer electrolyte fuel cell. J Electro-chem Soc; 142(1). p. 1–15

[7] Mosaad, M.I.; Ramadan, H.S. Power quality enhancement of grid-connected fuel cell using evolutionary computing techniques. *Int. J. Hydrog. Energy* 2018, 43, 11568-11582.

[8] Priya, K.; Sathishkumar, K.; Rajasekar, N. A comprehensive review on parameter estimation techniques for Proton Exchange Membrane fuel cell modelling. *Renew. Sust. Energ. Rev.* 2018, 93, 121-144.

[9] Hou, Y.; Zhou, B.; Zhou, W.; Shen, C.; He, Y. An investigation of characteristic parasmeter variations of the polarization curve of a proton exchange membrane fuel cell stack under strengthened road vibrating conditions. *Int. J. Hydrog. Energy* 2012, 37, 11887-11893.

[10] Salva, J.A.; Iranzo, A.; Rosa, F.; Tapia, E.; Lopez, E.; Isorna, F. Optimization of a PEM fuel cell operating conditions: Obtaining the maximum performance polarization curve. *Int. J. Hydrog. Energy* 2016, 41, 19713-19723.

[11] Selyari, T.; Ghoreyshi, A.A.; Shakeri, M.; Najafpour, G.D.; Jafary, T. Measurement of polarization curve and development of a unique semi-empirical model for description of PEMFC and DMFC performances. *CI & CEQ.* 2011, 17, 207-214.

[12] Zheng, Z.; Luo, L.; Zhu, F.; Cheng, X.; Yang, F.; Shen, S.; Wei, G.; Zhang, J. Degradation of core-shell Pt3Co catalysts in proton exchange membrane fuel cells (PEMFCs) studied by mathematical modelling. *Electrochimi. Acta* 2019, 323, 134751.

[13] Zerhouni M'hamed Houari, Zerhouni Fatima Zohra, zegrar Mansour. Photovoltaic Solar Array: Modeling and output power optimization [J]. Environmental Progress and Sustainable Energy,2016,10(1002): 1529-1536.

[14] Shubhankar Niranjan Deshkar, Sumedh Bhaskar Dhale, Jishnu Shekar Mukherjee. Solar PV array reconfiguration under partial shading conditions for maximum power extraction using genetic algorithm[J].Renewable and Sustainable Energy Reviews, 2015,4(3):102-110.

[15] Thomas Ott, Thomas Walter, Dimitrios Hariskos. Accelerated Aging and Contact Degradation of CIGS Solar Cells[J]. JOURNAL OF PHOTOVOLTAICS, 2013, 3(1): 514-518.

[16] Priya Ranjan Satpathy, Sasmita Jena, Renu Sharma. Power enhancement from partially shaded modules of solar PV arrays through various interconnections among modules[J]. Energy, 2018, 144: 839-850.

[17] Daniel Gomez Lorente, Simone Pedrazzi, Gabriele Zini. Mismatch losses in PV power plants[J]. SOLAR ENERGY, 2014, 42-49.

[18] Ye X, Zhao Y H, Quan Z H, Thermal management system of lithium-ion battery module based on micro heat pipe array[J]. ENERGY RESEARCH, 2017 : 648-655.

[19] J.Siecker, K. Kusakana, B.P.Numbi. A review of solar photovoltaic systems cooling technologies[J]. Renewable and Sustainable Energy Reviews, 2017, 79: 192-203.

[20] Wang Y.P, Shi X.S, Huang Q W. Experimental study on direct-contact liquid film cooling simulated dense-array solar cells in high concentrating photovoltaic system[J]. Energy Conversion and Management, 2017, 135: 55-62.

[21] Yang Peixin, ZHAO Xiqing. Ideal power supply in circuit theory [J]. Chinese Journal of Physics, 2005 (12) : 26-27

[22] Ding Jinlei. Principle Research and Application of Explicit Solution of I-V Equation for Solar Cell [D]. Ph.D thesis. Hefei: University of Science and Technology of China, 2007.

[23] Elodie Lechartier, Elie Laffly, Marie-Cécile Péra, Rafael Gouriveau, Daniel Hissel, NoureddineZerhouni. Proton Exchange Membrane Fuel Cell Behavioural Model Suitable for Prognositics[J]. International Journal of Hydrogen Energy 2015; 40: 8384-8397.

[24] E. Duran, J.M. Andujar, F. Segura. A high-flexibility DC load for fuel cell and solar arrays power sources based on DC-DC converters[J]. Applied Energy, 2011,88:1690-1702.

[25] Xian-liang, XIAO Jian, JIA Junbo. Experimental Study on Alternating Impedance Spectra of Proton Exchange Membrane Fuel Cells [J]. Chinese Journal of Electrical Engineering, 2010, 30(35): 101-106.

[26] Zhang Zhongzheng. Research on Solar Cell Application Theory. Doctoral dissertation [D]. Hefei: University of Science and Technology of China, 2014.6-12.

[27] Yuan Qibin. Theoretical Study on Power Mismatch of Battery Array. Master's thesis [D]. Hefei: University of Science and Technology of China, 2018.

[28] Junbom Kim, Seong-Min Lee, Supramaniam Srinivasan. Modeling of Proton Exchange Membrane Fuel Cell Performance[J]. JOURNAL OF THE ELECTROCHEMICAL SOCIETY, 1995, 142(8): 2670-2674.

[29] Xu zhidan, huang haiyan, Bao cheng. Semi-empirical model of proton exchange membrane fuel cell(in Chinese) [J]. Chinese journal of mechanical engineering, 2007,43 (10) :126-131.

[30] Weijie Shen and Dennis K.J.Lin. A Conjugate Model for Dimensional Analysis[J]. TECHNOMETRICS, 2018,60(1), 79-89.

[31] Kari L, Seppo V. Effects of PV Array Layout Electrical Configuration and Geographic Orientation on Mismatch Losses Caused by Moving Clouds[J]. Solar Energy, 2017, 144: 548–555.

[32] Zhang jinhui, Pei pucheng. Experimental Study on Ohmic Impedance of Proton Exchange Membrane Fuel Cell [J]. Journal of tsinghua university, 2007,47 (2) : 228-231.

[33] Cao Dianxue,WU Yanzhuo. Department of Fuel Cell [M]. Beijing: Beihang University Press 2009:49.

[34] Tan xunqiong,wu zhengqiu,zhou ye.Integrated Modeling and Simulation of Solid Oxide Fuel Cells [J]. Chinese journal of electrical engineering, 2010,30 (17)

Attention-based Convolutional Neural Networks for Diesel Fuel System Fault Diagnosis

Yijing Xie*
Air and Missile Defense College
Air Force Engineering University
Xi'an, China
xyj-093129@163.com

Tianlin Niu
Air and Missile Defense College
Air Force Engineering University
Xi'an, China
ntl2368@163.com

Siyu Shao
Air and Missile Defense College
Air Force Engineering University
Xi'an, China
cathygx.sy@gmail.com

Yuwei Zhao
Air and Missile Defense College
Air Force Engineering University
Xi'an, China
ywzhao1987@126.com

Yuemeng Cheng
Air and Missile Defense College
Air Force Engineering University
Xi'an, China
57855671@qq.com

Abstract—Diesel fuel system is a significant part of the diesel engine, whose stable and reliable working state is the key guarantee for safety and efficiency of the whole system. It is essential to conduct health monitoring and fault diagnosis for the diesel fuel system based on intelligent technology. Recent years, deep learning has become an effective means to perform intelligent fault diagnosis for various mechanical system and convolution neural networks have achieved remarkable success in various applications. In order to achieve efficient and accurate fault diagnosis for diesel fuel system, an improved convolution neural network combined with attention mechanism is established in this paper. The additional attention coefficients are learned to provide importance bias for the extracted features so that the deep model can make the most of the key signature associated with the working condition and ignore the irrelevant intermediate features in the consideration of computational efficiency. Experiments have been carried out and the results have proved the model effectiveness and efficiency.

Keywords—diesel engine, diesel fuel system, fault diagnosis, convolutional neural network, attention

I. INTRODUCTION

Diesel engine is a complex thermodynamic machine, whose normal operation requires a reliable supply of diesel from the diesel fuel system. In order to obtain satisfactory power, economy and emission indicators of the diesel engine, the diesel fuel system is required to be of high quality (high pressure spray and diesel injection rules), quantity (precise control of oil quantity), time (start point and duration of injection) and other aspects to achieve a reasonable match with the whole machine. The diesel fuel system is the core component of the diesel engine, and various failures are unavoidable during the work process, whose failures account for a relatively large proportion of the diesel engine failure types. When fault occurs in the diesel fuel system, the normal power output of the diesel engine will be affected, or even causing a fatal failure of the whole system. Therefore, appropriate health monitoring and fault diagnosis strategy for diesel fuel system is essential for stable operation [1-4].

With the continuous development of intelligent diagnosis technology, more and more up-to-date research hotspot focuses on how to ensure stable operation of the diesel engine system, track system state changes in real time and reduce its failure rate [5-8]. Cai et al. proposed an improved architecture to realize feature extraction from vibration signals and achieve desired performance for diesel engine fault diagnosis [9]. Zhao et al. using local wave time-frequency

analysis to conduct researches on diesel engine piston wear problems [10]. Liu et al. utilized self-adaptive Wigner-Ville distribution, improved fast correlation-based filter and relevance vector machine to achieve accurate fault diagnosis for diesel engines [11]. Similar to traditional fault diagnosis approaches for mechanical system [12,13], the above researches mainly rely on hand-crafted features, which requires prior knowledge of the specific tasks. The quality and applicability of the extracted features play an important in the fault recognition where the proper features contribute to accurate prediction while useless features may lead to misclassification. In order to reduce artificial dependency, deep learning-based approaches are widely investigated in the applications of mechanical fault diagnosis. Deep neural networks have the automatic representations learning ability from raw data, and is trained to learn useful features through parameter updates. For the complex diesel engine system, deep architecture-based architecture is able to learn hierarchical representations where the abstract features help to realize accurate fault prediction.

Inspired by the above researches, this paper aim to build a diesel fuel system fault diagnosis framework using attention mechanism and convolutional neural networks to study and estimate common faults of diesel fuel system. Experimental results have proved that the proposed strategy is effective to predict diesel system failure.

II. METHODOLOGY

A. Basic Structure of Diesel Engine

Diesel engine is mainly composed of two major institutions and four major systems, namely the body and crank connecting rod mechanism, gas distribution mechanism, diesel fuel system, lubrication system, cooling system and starting system, shown as in Figure 1.

Among these, diesel fuel system is a significant part of the diesel engine, whose main function is to provide clean diesel to the combustion chamber of each cylinder regularly, quantitatively, and in sequence according to the working order and load size of the diesel engine, meeting all the requirements of the combustion theory. The required degree of atomization is mixed with the air in the cylinder to form a combustible mixture, which will burn on its own to realize the conversion of diesel chemical energy into thermal energy.

---

*Yijing Xie is the corresponding author. (e-mail: xyj-093129@163.com).
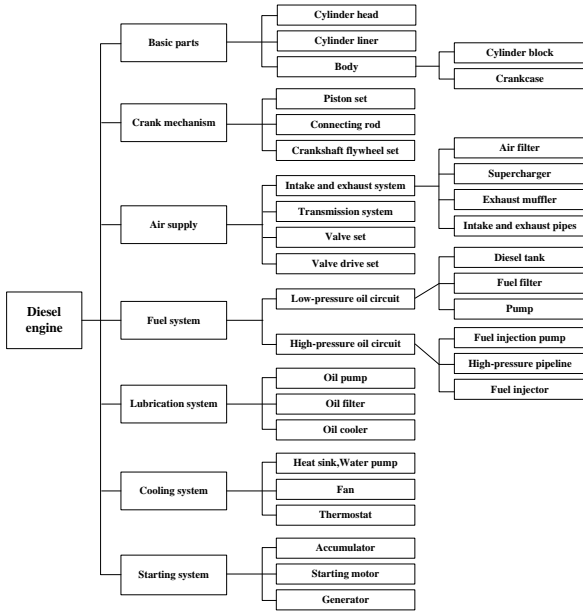
Fig. 1. Typical diesel engine system.

The complicated operating conditions of diesel engines require regular maintenance and inspection by users. Table I. shows the incidence of major failures of diesel engines.

TABLE I. MAJOR FAILURE TYPES IN DIESEL ENGINES

Fault Type	Fault Rate (%)	Fault Type	Fault Rate (%)
Diesel fuel system failure	27.0	Gear and drive failure	3.9
Water leakage fault	17.3	Governor gear failure	3.9
Valve and valve seat failure	11.9	Diesel leak	3.5
Bearing failure	7.0	Air leak	3.2
Piston assembly failure	6.6	Other system failure	2.5
Oil leakage and lubrication system failure	5.2	Base failure	0.9
Turbocharging system failure	4.4	Crankshaft failure	0.2

Based on the information above, diesel fuel system failure is the most common fault type during diesel system operation. Shown in Figure 2, the diesel fuel system is mainly composed of a diesel tank, a filter, a low-pressure diesel circuit, a diesel delivery pump, a high-pressure diesel circuit, a diesel injection pump, a diesel injector, a governor, etc. The complex system composition leads to an increased probability of failure.

The fault phenomenon is an important representation throughout the entire process of fault occurrence, and is also a significant basis for fault analysis and elimination. There are many kinds of diesel engine failures, and nearly one-third of the failures occur in the diesel system, which can be summarized in four main aspects: difficulty in staring, insufficient power, unstable speed and exhaust smoke.

The basic condition for the diesel engine to start is that the minimum speed of the start operation must be guaranteed. At the end of the compression, there must be a certain pressure and temperature. The cause of the failure of the

relevant diesel system is the lack of diesel in the diesel tank, air or water in the diesel fuel system, clogged diesel filter, clogged diesel line, diesel pump or intermittent pump oil, the diesel injector without diesel injection, low diesel injection volume or low diesel injection pressure, little or no pump oil in the injection pump and wrong the diesel fuel time. The cause of insufficient power is basically the same as the difficulty of starting.
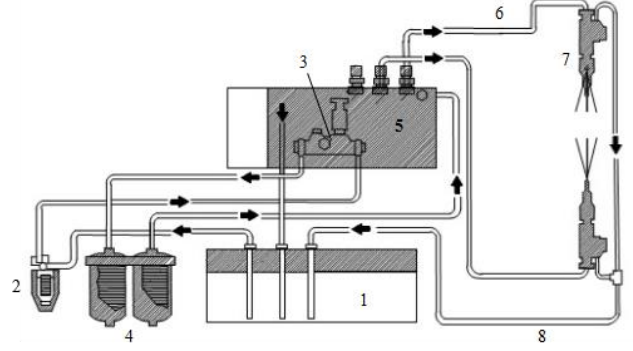
Fig. 2. Typical diesel fuel system.
1-Diesel tank；2-Diesel primary filter；3-Low pressure oil pump；4-Diesel fine filter；5-Fuel injection pump；6-High-pressure pipeline；7-Fuel injector；8-Return pipe

Based on the analysis of the common faults of diesel engines, the main causes of the diesel fuel system faults can be summarized as the presence of water or air in the oil path, blocked oil path, lack of oil or low amount of oil in the oil path, the injector with abnormal working conditions and the diesel injection pump with malfunction. These may influence the diesel fuel of the diesel system, which may inevitably lead to changes in the diesel fuel state, that is, the diesel flow rate, pressure and other parameters will also change eventually, resulting in the changes in the waveform state and parameter values on the pressure wave of the high-pressure diesel pipe. Therefore, through the analysis and research of the pressure waveform of the diesel engine system, the representative characteristic parameters are extracted. Based on efficient an effective data processing, the working state of the diesel system can be estimated exactly. Thus, current working condition of the diesel engine can be judged and the fault diagnosis of the diesel system can be realized.

When the state of a certain part of the diesel system changes, it will cause distortion of the diesel pressure waveform. As long as the characteristics of the waveform distortion are grasped, the distortion waveform is analyzed and diagnosed. The fault diagnosis process of the diesel system can be transformed into a pattern recognition issue based on high-pressure diesel pip pressure waveform analysis. According to the existing experimental data, as to the morphological structure of the pressure waveform, the state signals such as maximum pressure, seating pressure, ejection pressure, sub-maximum pressure, waveform amplitude, rising edge width, waveform width and maximum residual wave width, are able to reflect the working conditions of the diesel fuel system. Therefore, the fault diagnosis framework based on diesel pressure waveform analysis is typical and effective.

B. Convolutional Neural Network

As one basic building block in deep learning architecture, multiple convolution layers are stacked to form the

Convolutional Neural Network (CNN), which is one of the most successful deep model in the real world applications. CNNs have superior feature learning ability [14] and have achieved various success in the field of computer vision [15], natural language processing [16], speech recognition and biology [17]. The main concepts used in CNN such as Local Receptive Fields, Shared Weights and Pooling, help reduce the number of parameters that need to be trained, therefor helping increase the training efficiency of the neural network [18].

CNN contains three main building blocks: convolution layer, pooling layer and fully-connected layer. Convolutional layers serve as feature extractor to scan across the input and generate various feature maps according to different convolutional kernels, where each convolutional kernel learns one certain kind of features. Mathematically, convolution operation can be calculated as:

$$a_{i,j} = a(\sum_{d=0}^{D-1}\sum_{m=0}^{F-1}\sum_{n=0}^{F-1} w_{d,m,n} x_{d,i+m,j+n} + b) \tag{1}$$

where input x has the depth of D and $x_{d,i+m,j+n}$ represents the area (i, j) within d-th depth of the input; $F \times F$ denotes the size of convolutional kernel, $a_{i,j}$ denotes the area (i, j) of the output feature map; weights connected to the d-th depth convolutional kernel are shown as $w_{d,m,n}$.

For further improving the training efficiency of deep model, pooling operation is conducted on the extracted feature maps. Detailed, pooling operation is calculated as:

$$y_{i,j,k} = pool_{(m,n)\in R_{ij}}(x_{m,n,k}) \tag{2}$$

Where the pooling function mainly contains two types: average pooling and max pooling, denotes as $pool(\cdot)$.

$y_{i,j,k}$ denotes the pooling operation output, while $x_{m,n,k}$ is the input area.

Generally, convolutional layers combined with pooling layers are defined as a basic building block of the deep architecture, where one pooling layer is connected to one or more convolutional layers. Several building blocks are stacked to form a deep CNN architecture to obtain the hierarchical presentation learning ability. As the model depth increases, the representations from upper hidden layers become more abstract, which is helpful to achieve accurate prediction [19, 20].

After certain number of stacked building blocks, the deep architecture has the capability to learn useful representations for recognition tasks. Finally, fully-connected layers are designed to perform prediction [21-23]. Mathematically, fully-connected layer can be defined as:

$$y = f(Wa + b_c) \tag{3}$$

Where a is the extracted feature maps from the building blocks and the fully-connected layer generates the possible label y as the model output.

C. Attention Mechanism

Attention mechanism is a strategy to add importance bias to the deep architecture for the purpose of making full use of the key features, which helps solve the problem of information overload when the extracted features are enormous using deep neural network. For a deep architecture, large numbers of parameters usually means strong representation learning ability with mass information stored. However, the large amount information may contain redundant items which have no contributions to accurate classification and even causing heavy training burdens. Therefore, attention coefficients are utilized to pick the most important representations from the neural network and ignore the useless information, which may improve the model training efficiency.

Mathematically, attention coefficients can be calculated from:

$$c_i = \sum_{j=1}^{n} \alpha_{ij} h_j \tag{4}$$

$$\alpha_{ij} = \frac{\exp(s(h_t, \overline{h_s}))}{\sum_{k=1}^{n} \exp(s(h_t, \overline{h_s}))} \tag{5}$$

$$s(h_t, \overline{h_s}) = h_t \overline{h_s} \tag{6}$$

Where c_i is the weighted average of the intermediate convolutional layer, and α denotes the weight scales between the adjacent layers. s represents the weight calculation method. Attention mechanism aims to selectively leverage the possible intermediate representations, and construct the mapping relationship between hidden features and the final output using corresponding coefficients [24]. Compared with traditional CNN, attention mechanism is able to reduce the whole number of model parameters, therefore improving the model efficiency.

III. FAULT DIAGNOSIS FRAMEWORK

A fault diagnosis framework based on attention mechanism and convolution neural networks is designed for diesel fuel system fault diagnosis. Attention-based model is able to learn representations with emphasis and CNN model is able to extractor effective fault signatures for accurate fault classification. With the unique structure of attention-based convolutional neural network (ATT-CNN), accurate and efficient fault diagnosis for diesel fuel system can be achieved. The whole framework for diesel fuel system fault diagnosis is shown in Figure 3.

The whole procedure of the proposed pipeline can be divided into 3 main stages.

- Data acquisition: based on the information from Section I, the pressure wave of the high-pressure diesel pipe is an effective indicator to reflect working condition of the diesel fuel system. Therefore, the fuel pressure is collected as the training data of the proposed fault diagnosis framework. The original collected pressure data serve as the input data and for model training, collected data is divided into different sample slides for training and testing separately.

- Feature extraction: three-layers convolutional neural network is designed as the main structure of the

proposed approach. Besides, attention module is designed afterwards for emphasizing features with importance in accurate working condition prediction. CNN with attention coefficients is able to achieve faster training speed and higher classification accuracy eventually.
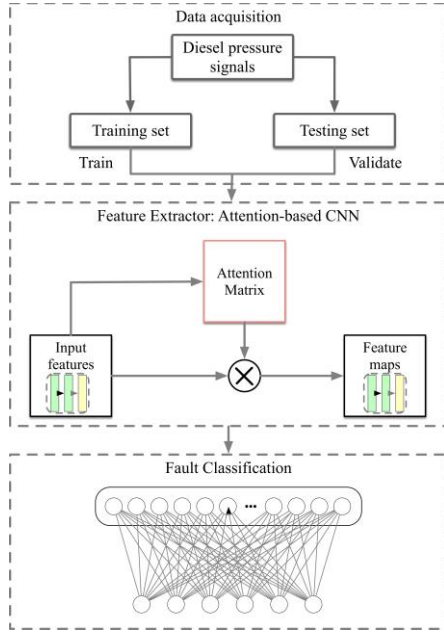
Fig. 3.   Diesel fuel system fault diagnosis framework.

- Fault classification: the fully-connected layer is placed after the feature extractor block to realize label prediction. The error between predicted labels and the true labels is calculated and the model parameters are updated through error back propagation.

IV.   EXPERIMENTAL VERIFICATION

For model verification, several experiments have been carried out. As for the dataset preparation for training and testing, collected pressure data serve as the original input of the proposed fault diagnosis framework and are divided into training and testing subsets, with 200 training samples and 100 testing samples. In order to assess the performance of the proposed architecture, classification accuracy for fault category is calculated, defined as:

$$accuracy = \frac{n_{correct}}{N_{total}} \tag{7}$$

where $n_{correct}$ is the number of samples that the proposed model predicts with correct labels while N_{total} represents the number of total samples. Classification accuracy is the correctly predicted percentage, which indicates how accurate prediction the designed model is able to achieve in diesel fuel system failure recognition tasks.

The data used in the experiments are collected from 4135 diesel engine, which contains 1 normal working condition and 4 different faults, including stucked needle valve, leakage in needle valve, delivery valve failure and lack of oil supply. Therefore, the diesel fuel system fault diagnosis based on the current dataset can be treated as a 5-class recognition issue.

For comparison, additional experiments have been conducted to explore the model performances in failure type prediction using various machine learning frameworks. The used comparison architectures are listed as follows:

- 3-layer back propagation neural network using unprocessed sensor data (BP-NN)

- 2-layer back propagation neural network based on Hand-crafted features extracted in advance

- 3-layer Deep Belief Network using unprocessed sensor data (DBN+BP)

- 6-layer Deep Neural Network using unprocessed sensor data (DNN)

- Convolutional Neural Network without attention mechanism using original signals (CNN)

Training and testing time based on 50 repeated experiments are recorded to investigate the model efficiency, and in order to obtain a robust estimate of proposed method, 10-fold cross validation is carried out and statistics are shown to summarize the model performances. Results are shown in Table II and figure 4.

TABLE II.   FAULT DIAGNOSIS RESULTS BASED ON VARIOUS METHODS

Methods	Classification Results	
	Accuracy	Training time
BP-NN	90.08% (4.75)	12 epochs/155s
Hand-crafted features+BP	98.56% (1.67)	5 epochs/69s
DBN+BP	95.65% (0.75)	8 epochs/121s
DNN	98.44% (1.05)	20 epochs/225s
CNN	97.32% (0.67)	7 epochs/101s
ATT-CNN	**98.75%**(0.25)	**4 epochs/66s**
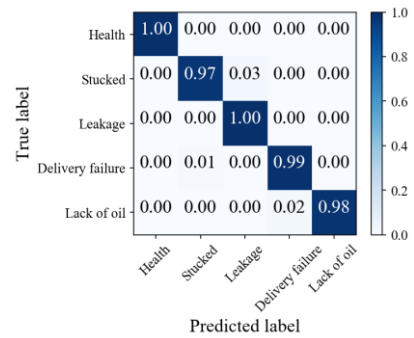
Fig. 4.   Confusion matrix of the ATT-CNN model.

Based on the experimental results above, the proposed method outperforms other machine learning-based methods in both efficiency and accuracy. From the results, hand-crafted features based on prior knowledge are well-selected in advance for the specific fault recognition task so that the final classification accuracy and model training time is desirable. Compared with CNN without attention coefficients, the proposed ATT-CNN has higher classification accuracy and less training time, which indicates the attention mechanism is able to boost the performance of deep architecture. The added importance bias helps to select useful intermediate features for fault

classification and ignore useless representations to save model training time so that both model effectiveness and model efficiency can be achieved.

V. Conclusion

This paper proposes an attention-based convolutional neural network framework for diesel fuel system fault diagnosis. Compared with traditional CNN model, attention mechanism is designed to learn the degree of importance of different extracted features and therefore making the most of the useful representations to improve the model prediction accuracy and save the model training time. Experiments among diesel pressure data have proved the effectiveness of the proposed fault diagnosis pipeline.

Future study can be conducted to improve the performance of the attention-based model and search for an optimal strategy of combination.

Acknowledgment

References

[1] N. B. Jones and Y.H. Li, "A review of condition monitoring and fault diagnosis for diesel engines," Tribotest, vol.6, no.3, pp. 267–291, 2000.

[2] L.A. Malm, A. Hultman, J. Enstrom, Main Engine Damage, Swedish Club, 2015.

[3] R.Y. Jiang, X.P. Yan, "Condition monitoring of diesel engines," in: Complex System Maintenance Handbook, 2008, pp. 533–557.

[4] X.P. Yan, X.J. Xu, C.X. Sheng, et al., "Intelligent wear mode identification system for marine diesel engines based on multi-level belief rule base methodology," Meas. Sci. Technolk, vol. 29, no. 1, pp. 1–13, 2018.

[5] X.P. Yan, Condition Monitoring and Fault Diagnosis for Mechanical System, Wuhan University of Technology Press, Wuhan, 2009.

[6] F. Zhang, J. Huang, F. Chu, and L. Cui, "Mechanism and method for outer raceway defect localization of ball bearings," IEEE Access, vol. 8, pp. 4351–4360, 2020.

[7] X. Xu, Z. Zhao, X. Xu et al., "Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models," Knowledge-Based Systems, vol. 190, Article ID 105324, 2019.

[8] C. Wang, Y. Zhang, and Z. Zhong, "Fault diagnosis for diesel valve trains based on time-frequency images," Mechanical Systems and Signal Processing, vol. 22, no. 8, pp. 1981–1993, 2008.

[9] Y. Cai, G Xu, A Li and X Wang, "A Novel Improved Local Binary Pattern and Its Application to the Fault Diagnosis of Diesel Engine," Shock and Vibration, vol. 2020, Article ID 9830162, 2020.

[10] F. Zhao, G. Li, H. Hu et al., "Research on diesel engine piston wear fault diagnosis method based on the local wave time-frequency analysis," The Open Automation and Control Systems Journal, vol. 6, no. 1, pp. 913–918, 2014.

[11] Y. Liu, J. Zhang, and L. Ma, "A fault diagnosis approach for diesel engines based on self-adaptive WVD, improved FCBF and PECOC-RVM," Neurocomputing, vol. 177, pp. 600–611, 2016.

[12] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques-part I: fault diagnosis with model-based and signal-based approaches," IEEE Transactions on Industrial Electronics, vol. 62, no. 6, pp. 3757–3767, 2015.

[13] Q. He, X. Wang, and Q. Zhou, "Vibration sensor data denoising using a time-frequency manifold for machinery fault diagnosis," Sensors, vol. 14, no. 1, pp. 382–402, 2013.

[14] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis," IEEE Transactions on Instrumentation and Measurement, vol. 66, no. 8, pp. 1926–1935, 2017.

[15] H. Azizpour, A. Sharif Razavian, J. Sullivan, A. Maki, and S. Carlsson, "From generic to specific deep representations for visual recognition," in Proc. IEEE Conf. Comput. Vision Pattern Recognit. Workshops, 2015, pp. 36–45.

[16] R. Socher, Y. Bengio, and C. D. Manning, "Deep learning for NLP (without magic)," in Tutorial Abstracts of ACL 2012. Stroudsburg, PA, USA: Association for Computational Linguistics, 2012, pp. 5–5.

[17] P. Baldi, "Deep learning in biomedical data science," Annu. Rev. Biomed. Data Sci., vol. 1, pp. 181–205, 2018.

[18] S. Shao, R. Yan , Y. Lu, et al. "DCNN-based multi-signal machine fault diagnosis," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 6, pp. 2658-2669, 2020.

[19] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R.Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," 2012, arXiv:1207.0580.

[20] J. Gu et al., "Recent advances in convolutional neural networks," 2015, arXiv:1512.07108.

[21] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, and G. Wang, "Recent advances in convolutional neural networks," arXiv preprint arXiv:1512.07108, 2015.

[22] A. Karpathy, "Cs231n convolutional neural networks for visual recognition," http://http://cs231n.github.io/convolutional-networks/.

[23] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in European conference on computer vision. Springer, 2014, pp. 818–833.

[24] L. Wu, Y. Wang, X. Li et al, "Deep attention-based spatially recursive networks for fine-grained visual recognition," IEEE T. Cybern., vol. 49, no. 5, pp. 1791–1802, Mar.2018.

RUL Prediction For Turbine Disc Based On High-Order Particle Filtering

Yang Fu
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
yangfu@stu.xjtu.edu.cn

Hongrui Cao*
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
chr@mail.xjtu.edu.cn

Abstract—Turbine disc is one of the most crucial parts in an aero-engine and it is also regarded as the vulnerable part which suffers from damage frequently. Remaining useful life (RUL) prediction technology is a useful tool to provide failure warnings and increase safety. For turbine disc, however, it is still a great challenge. This paper proposes a high-order particle filtering (HOPF)-based approach for RUL prediction of turbine disc. First, a turbine disc health indicator is constructed based on the unbalance vibration analysis. Then, an improved double exponential model is developed for turbine disc degradation modeling. Next, model updating and RUL prediction are carried out by the HOPF algorithm. The proposed approach is finally validated using the experimental data of a turbine disc. Satisfactory results demonstrate that the proposed approach performs well on turbine disc RUL prediction.

Keywords—turbine disc, remaining useful life, degradation modeling, high-order particle filtering

I. INTRODUCTION

As one of the most crucial parts, turbine disc is used for assembling blades and transferring power. Suffering from the centrifugal loads and thermal stresses, turbine disc is regarded as a vulnerable part [1]. Failures of turbine discs influence the functionality of engines, and then cause catastrophic consequences. Such failures are generally intolerable. In order to avoid that, much effort has been focused on predicting the design life for turbine disc by using damage mechanics and reliability theory [2-3]. Then, scheduled maintenance tasks are arranged based on the predicted design life. However, the scheduled maintenance strategy for turbine disc could produce some problems. (i) Many turbine discs are early replaced. (ii) A minority of turbine discs fail before maintenance. (iii) The expense of maintenance task is huge since the waste of potential life, and the cost of removing and assembling the aero-engine. In order to avoid that, remaining useful life (RUL) prediction has emerged as a critical technology to prevent the turbine disc from sudden failure and optimize maintenance schedules, which is meaningful to reduce maintenance waste and guarantee flight safety [4].

Many works have been focused on RUL prediction for machinery. The major approaches could be divided into three types: model-based, data-driven, and data-model fusion. Model-based approaches rely on physical or experiential failure mechanism for predicting the machinery RUL [5]. However, it is generally difficult to construct high fidelity failure model for a complex machinery system. Data-driven approaches use large amounts of measured data to model the degradation trend of machinery and predict the RUL by using machine learning techniques [6]. The performance of data-driven approaches depend on the quantity and quality of the data. Nevertheless, gathering enough eligible data for turbine

disc is very hard in practice since it is a time-consuming and costly task. Data-model fusion approaches have been the hotspot of RUL prediction in recent years since they can take full advantages of experiential knowledge from failure model and the real-time information from the monitoring data [7]. Among them, PF-based data-model fusion approaches have been proven to be an efficient and effective way since it is capable to track and predict the non-linear non-Gaussian system. In these approaches, the physical/data/empirical model is treated as the prior knowledge of machinery degradation characteristic to construct a state-space model, and real-time monitoring data is treated as the likelihood information form an individual degradation process. Based on the Bayesian theory, the model and real-time data can be fused into posteriori knowledge about degradation state with the consideration of the influence of uncertainty. With the help of the Monte Carlo technology, PF shows a strong capacity on individual degradation tracking and RUL prediction. For example, Peng et al. [8] combined the deep belief network with an improved PF and utilized this model for RUL prediction of aircraft engine. Yang et al. [9] described the performance degradation by the exponential model and employed PF for forecasting the RUL of an auxiliary power unit in the aircraft.

In the context of turbine disc RUL prediction, however, two problems still exist in PF-based approaches. The first problem is how to construct an effective health indicator (HI) from turbine disc monitoring data. A good HI with an obvious monotonic trend can simplify the modeling process and produce accurate predictions. The second problem is that most PF-based approaches are limited to the first-order Markov assumption. It considers that the current degradation state only relies on the previous degradation state. In the real case of turbine disc, however, this assumption is not true in reality. The high-order Markov assumption may be more suitable for depicting the degradation process in which the current degradation state depends not only on the previous state but also on multiple *m*-step-before states.

In order to deal with the two abovementioned issues, we present a high-order particle filtering (HOPF)-based approach for RUL prediction of turbine disc. Our contributions mainly include the following aspects.

(1) An HI is constructed for turbine disc based on the unbalance vibration analysis.

(2) An improved double exponential model is developed for turbine disc degradation modeling, and RUL is predicted by the HOPF based algorithm.

(3) A low cycle fatigue test of turbine disc is implemented to validate the proposed approach.

The remainder is summarized as follows. The theoretical foundation of HOPF is given in Section II. Section III presents the HOPF-based RUL prediction approach for turbine disc. Section IV conducts an experimental validation. Finally, conclusion is summarized in the last section.

II. HIGH-ORDER PARICEL FILTERING

The state-space model with an high-order Markov assumption is expressed as follows:

$$\begin{cases} x_k = f_k\left(x_{k-1}, x_{k-2}, \text{L } x_{k-m}, \omega_{k-1}\right) \\ z_k = h_k\left(x_k, v_k\right) \end{cases} \quad (1)$$

where x_k denotes the state at time t_k, f_k is the state transition equation. ω_{k-1} denotes the process noise at time t_{k-1}. z_k is the measurement at time t_k, h_k denotes the measurement equation, v_k denotes the measurement noise.

The prior probability density function (PDF) at time t_k can be calculated by the prediction step:

$$p\left(x_{0:k}\left|z_{1:k-1}\right.\right) = \int p\left(x_k\left|x_{k-m:k-1}\right.\right)p\left(x_{0:k-1}\left|z_{1:k-1}\right.\right)\mathrm{d}x_{0:k-1} \quad (2)$$

where $p(x_k|x_{k-m:k-1})$ is the probabilistic state model defined in $f_k(\cdot)$, $p(x_{0:k-1}|z_{1:k-1})$ is the state PDF at time t_{k-1}. With the assumption of mth-order Markov property, $p(x_k|x_{0:k-1}, z_{1:k-1}) = p(x_k|x_{k-m:k-1})$

The posterior PDF of state can be obtained by the update step when a new measurement z_k is available:

$$\begin{aligned} p\left(x_{0:k}\left|z_{1:k}\right.\right) &= \frac{p\left(z_k\left|x_k\right.\right)p\left(x_{0:k}\left|z_{1:k-1}\right.\right)}{p\left(z_k\left|z_{1:k-1}\right.\right)} \\ &= \frac{p\left(z_k\left|x_k\right.\right)p\left(x_k\left|x_{k-m:k-1}\right.\right)p\left(x_{0:k-1}\left|z_{1:k-1}\right.\right)}{p\left(z_k\left|z_{1:k-1}\right.\right)} \\ &\propto p\left(z_k\left|x_k\right.\right)p\left(x_k\left|x_{k-m:k-1}\right.\right)p\left(x_{0:k-1}\left|z_{1:k-1}\right.\right) \end{aligned} \quad (3)$$

where $p(z_k|x_k)$ denotes the likelihood function, $p(z_k|z_{1:k-1})$ denotes the normalizing item, which can be calculated as follows:

$$p\left(z_k\left|z_{1:k-1}\right.\right) = \int p\left(z_k\left|x_k\right.\right)p\left(x_k\left|z_{1:k-1}\right.\right)\mathrm{d}x_k \quad (4)$$

HOPF employs a series of particles { $x_{0:k}^i$, $i = 1, 2, \ldots, Ns$} with associated weights { w_k^i, $i = 1, 2, \ldots, Ns$} to approximate the posterior PDF, shown as:

$$p\left(x_{0:k}\left|z_{1:k}\right.\right) \approx \sum_{i=1}^{Ns} w_k^i \delta\left(x_{0:k} - x_{0:k}^i\right) \quad (5)$$

where Ns is the particle number, $\delta(\cdot)$ denotes the Dirac delta function.

If $x_{0:k}^i$ are drawn from an importance distribution $q(x_{0:k}|z_{1:k})$, the normalized weights are updated by the following equation:

$$w_k^i \propto \frac{p\left(x_{0:k}^i\left|z_{1:k}\right.\right)}{q\left(x_{0:k}^i\left|z_{1:k}\right.\right)}, \quad w_k^i = w_k^i \bigg/ \sum_{i=1}^{Ns} w_k^i \quad (6)$$

For simplifying the calculation, $q(x_{0:k}|z_{1:k})$ can be factorized as follows:

$$\begin{aligned} q\left(x_{0:k}\left|z_{1:k}\right.\right) &= q\left(x_k\left|x_{0:k-1}, z_{1:k}\right.\right)q\left(x_{0:k-1}\left|z_{1:k-1}\right.\right) \\ &= q\left(x_k\left|x_{k-m:k-1}, z_k\right.\right)q\left(x_{0:k-1}\left|z_{1:k-1}\right.\right) \end{aligned} \quad (7)$$

Then, particle weights can be expressed in the recursive equation:

$$w_k^i \propto w_{k-1}^i \frac{p\left(z_k\left|x_k^i\right.\right)p\left(x_k^i\left|x_{k-m:k-1}^i\right.\right)}{q\left(x_k^i\left|x_{k-m:k-1}^i, z_k\right.\right)}, \quad w_k^i = w_k^i \bigg/ \sum_{i=1}^{Ns} w_k^i \quad (8)$$

Here, the importance distribution $q(x_k|x_{k-m:k-1},z_k)$ is chosen as $p(x_k|x_{k-m:k-1})$, (8) can be simplified as follows:

$$w_k^i \propto w_{k-1}^i p\left(z_k\left|x_k^i\right.\right), \quad w_k^i = w_k^i \bigg/ \sum_{i=1}^{Ns} w_k^i \quad (9)$$

III. PROPOSED METHOD

Fig. 1 plots the framework of the proposed turbine disc RUL prediction approach. It consists of two main modules, i.e., HI construction for turbine disc, and HOPF-based RUL prediction. First, an HI is constructed for turbine disc based on the unbalance vibration analysis. Then, an improved double exponential model is developed to construct an high-order state-space model (HOSSM) for describing the degradation process of turbine disc. Next, the parameters and state are updated by HOPF. Finally, the RUL are predicted by a multi-step ahead prediction algorithm. More details are discussed as follows.
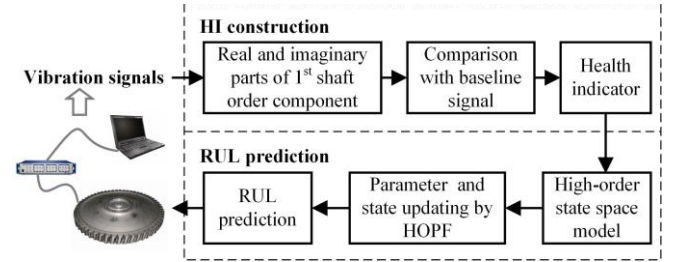
Fig. 1. The flowchart of the proposed approach.

A. HI Construction for Turbine disc

Fatigue crack is one of the most primary failure modes in a turbine disc. The crack in a turbine disc introduces an extra mass unbalance which produces unique vibration responses. Here, the first shaft order vibration is used for constructing the HI for turbine disc since it is often taken as a powerful feature of rotating unbalance [10]. Generally, the mass unbalance caused by the crack and the initial unbalance make up the unbalance of the turbine disc. However, only the unbalance caused by a crack is contributed to the degradation modeling and RUL prediction. Therefore, in order to extract the crack unbalance vibration characteristics, an HI is constructed as follows:

$$HI_k = \frac{1}{\sqrt{\left(R_k - R^*\right)^2 + \left(I_k - I^*\right)^2}} \quad (10)$$

where R_k and I_k are the real part and the imaginary part of the first shaft order component at time t_k (cycle k), respectively. R^* and I^* are respectively the real part and the imaginary part of the first shaft order component of initial mass unbalance. R_k and I_k are treated as the monolithic unbalance characteristic, R^* and I^* are treated as the initial mass unbalance characteristic. In the condition monitoring of turbine disc, the HI at a fixed

speed is derived by interpolation in each fatigue cycle (excursions from low speed to high speed back to low speed). The longer crack means the lower HI. Then, the changes of the HI can be modeled for health monitoring and RUL prediction of turbine disc.

B. High-order state-space model

The double exponential model which well balances the accuracy and complexity of modeling is widely used in degradation modeling [11]. The original model is expressed as follows:

$$x_k = a \exp(bt_k) + c \exp(dt_k) \tag{11}$$

where x_k denotes the current degradation state. a, b, c, and d are four model parameters which need to be evaluated according to the degradation data.

Then, the original double exponential model can be rewritten as a state-dependent form with second-order Markov property:

$$x_k = x_{k-1} \exp(b\Delta t) - x_{k-2} \exp(d\Delta t) \tag{12}$$

where $\Delta t = t_k - t_{k-1}$.

We use turbine disc HI data for fitting analysis of the original model, and find that the original double exponential model exists a shortcoming about amplitude bias, which means the modeling performance of the double exponential model is seriously influenced by the absolute amplitude of HI data. Therefore, a new model parameter named amplitude bias ε is introduced into the original model to overcome above-mentioned shortcoming. The improved double exponential model is denoted as follows:

$$x_k = (x_{k-1} - \varepsilon) \exp(b\Delta t) - (x_{k-2} - \varepsilon) \exp(d\Delta t) + \varepsilon \tag{13}$$

With the consideration of the uncertainty of process and the measurement, the HOSSM can be expressed as follows:

$$\begin{cases} x_k = (x_{k-1} - \varepsilon) \exp(b\Delta t) - (x_{k-2} - \varepsilon) \exp(d\Delta t) + \varepsilon + \omega_k \\ s_k = x_k + v_k \end{cases} \tag{14}$$

where x_k denotes the current degradation state of turbine disc, which is hard to be obtained and is represented by HI. s_k is the measured HI with the measurement noise v_k: $N(0, \sigma^2)$. The process noise ω_k can be ignored in most cases since it can be represented by the parameter uncertainty.

C. HOPF-Based parameter Updating and State Evaluation

The unknown parameters are denoted as a parameter vector $\boldsymbol{\Theta} = (b, d, \sigma, \varepsilon)'$. To simultaneously update the degradation state of turbine disc and model parameters, both degradation state x and model parameter $\boldsymbol{\Theta}$ are treated as the system state variable of HOPF, i.e., $\{z_0^i\}_{i=1:Ns} = \{x_0^i; \boldsymbol{\Theta}_0^i\}_{i=1:Ns}$. The particle weights are assigned as $w_0^i = 1/Ns$. Then, new particles $\{z_k^i\}_{i=1:Ns}$ can be obtained by:

$$z_k^i = \begin{bmatrix} (x_{k-1}^i - \varepsilon_k^i) \exp(b_k^i \Delta t) - (x_{k-2}^i - \varepsilon_k^i) \exp(d_k^i \Delta t) + \varepsilon_k^i \\ \boldsymbol{\Theta}_{k-1}^i \end{bmatrix} \tag{15}$$

According to (6) - (9), the new measured HI s_k at time t_k can be used to update and normalize particle weights as follows:

$$w_k^i = w_{k-1}^i \frac{1}{\sqrt{2\pi}\sigma_k^i} \exp\left(-\frac{(s_k - x_k^i)^2}{2(\sigma_k^i)^2}\right), \quad w_k^i = w_k^i \bigg/ \sum_{i=1}^{Ns} w_k^i \tag{16}$$

A resampling step is implement here to avoid particle degeneracy, and all particle weights are reset to $1/Ns$.

Then, the updated PDFs of degradation state and model parameters at time t_k can be obtained by:

$$p(z_k | s_{0:k}) \approx \sum_{i=1}^{Ns} w_k^i \delta(z_k - z_k^i) \tag{17}$$

D. RUL Prediction

The l-step prediction for turbine disc degradation state is obtained by:

$$p(x_{k+l} | s_{0:k}) \approx \sum_{i=1}^{Ns} w_k^i \delta(x_{k+l} - x_{k+l}^i) \tag{18}$$

$$x_{k+l}^i = (x_{k+l-1}^i - \varepsilon_k^i) \exp(b_k^i \Delta t) - (x_{k+l-2}^i - \varepsilon_k^i) \exp(d_k^i \Delta t) + \varepsilon_k^i \tag{19}$$

Then, the RUL predicted by the ith particle are expressed by:

$$RUL_k^i = \inf\{l : x_{k+l}^i \leq \tau | \boldsymbol{x}_{0:k}, \boldsymbol{\Theta}_k\} \tag{20}$$

where τ is the failure threshold.

Finally, the RUL PDF is approximately obtained by the following equation:

$$p(RUL_k | s_{0:k}) \approx \sum_{i=1}^{Ns} w_k^i \delta(RUL_k - RUL_k^i) \tag{21}$$

E. Implementation Steps

The detailed implementation process for the proposed approach are presented as follows.

Step 1: Acquire the turbine disc baseline vibration signal at a fixed speed in the un-cracked operation state.

Step 2: Acquire the real-time vibration signal at the same speed in each subsequent cycle and construct HI according to (10).

Step 3: Establish high-order state-space model by (14), and update model parameters and degradation state using (15) – (17) when a new measured HI of the turbine disc is available.

Step 4: Extrapolate the degradation trajectory of the turbine disc by (18) and (19), and predict the RUL PDF using (20) and (21).

Step 5: Repeat step 4 and step 5 when another new measured HI of the turbine disc is available.

IV. EXPERIMENTAL DEMONSTRATION

A. Experimental Setup

The low cycle fatigue test of the turbine disc is carried out on the Test Devices Inc. (TDI) spin test rig shown in Fig. 2. The maximum and minimum speeds are 15000rpm and 2000rpm, respectively. Each cycle takes 60 seconds. After 13927 cycles, the turbine disc is broken. Fig. 3 plots the overview of data acquisition. During the speed-up phase of each cycle, vibration displacement signals are collected at

different speeds, and the real and imaginary parts of the first shaft order vibration component are outputted by the TDI system. Fig. 4 plots the real part R and the imaginary part I of the first shaft order vibration component at 98% of the maximum speed after 13800 cycles.

Fig. 2. TDI spin test rig.
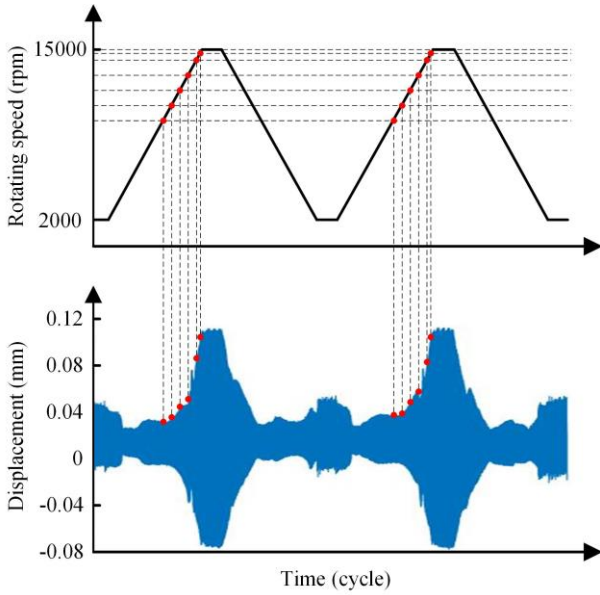
Fig. 3. Overview of data acquisition.
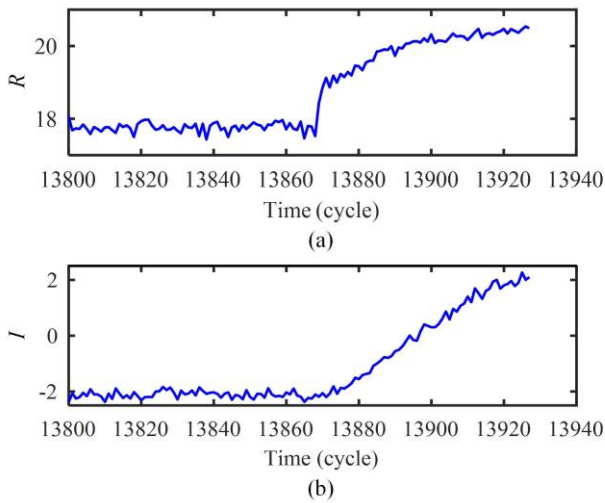
Fig. 4. Real part R and imaginary part I of the first shaft order vibration.

B. Results and Discussion

The data collected by the TDI system is utilized to construct the HI by using (1). Here, R^* and I^* are obtained by averaging the data gathered at first 100 cycles. To show the robustness of the proposed turbine disc HI, 6 HIs established at different speeds are plotted in Fig. 5. It can be seen that the proposed HI with a monotony evolution trajectory is good at representing the degradation process of turbine disc. In this section, data gathered at 98% of the maximum speed, i.e., 14700rpm is used to validate the proposed prediction approach. Here, the detection threshold is adaptively constructed by the 3σ interval. First, the historical HIs of turbine disc in the normal stage can be used to establish the bound $[\mu-3\sigma, \mu+3\sigma]$. To avoid the false alert caused by random noises, the degradation starting time is determined when 5 consecutive HIs passed through the 3σ interval. Fig. 6 plots the degradation detection for the turbine disc. It is seen that the turbine disc runs stably before 13880 cycles, and then the HI curve shows a decreasing trend, which means the turbine disc is degrading.
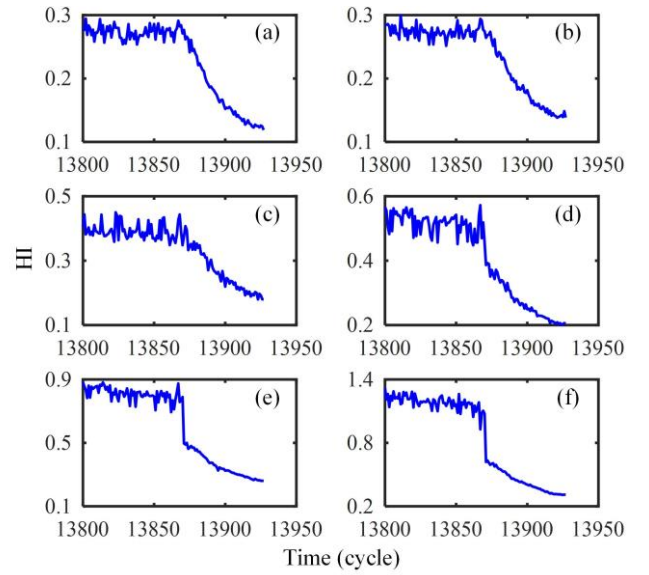
Fig. 5. Proposed HI for turbine disc condition monitoring at different speed. (a) 14700rpm, (b)14250rpm, (c) 13500rpm, (d) 12000rpm, (e) 10500rpm, and (f) 9000rpm.
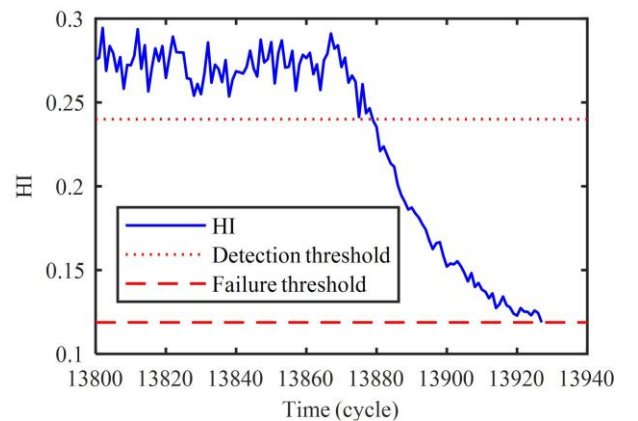
Fig. 6. Performance degradation of turbine disc.

In the high-order state-space model, initialization plays a pivotal role to obtain suitable model parameters. A good initialization can give accurate prediction results from an early

point because it can reduce the convergence time. In this case study, the first 25 HIs are used to initialize the high-order state-space model by the Metropolis-Hastings algorithm [12]. Then, the model parameter updating and RUL prediction are carried out according to the HIs from 13880 to 13927 cycles by the HOPF algorithm. The model parameter updating results including median and 95% confidence interval (95% CI) are shown in Fig. 7. It can be observed that the model parameters are dynamically updated in real time when a new HI is available. Moreover, the 95% CI of parameter evaluation converges to a narrow range as time goes on, which reflects that the uncertainty of model parameter is decreased.

Then, the degradation state of the turbine disc is evaluated by HOPF algorithm as shown in Fig. 8 (blue line marked by small circle). In order to show the advantage of the HOPF-approach (Approach C) comparing with standard PF, two classical approaches, i.e. the Paris model with standard PF [13] (Approach A), and the exponential model with standard PF [14] (Approach B), are also employed to evaluate the degradation state of the turbine disc. It is observed that the HOPF based approach catches measurements better and evolves smoother than the others.
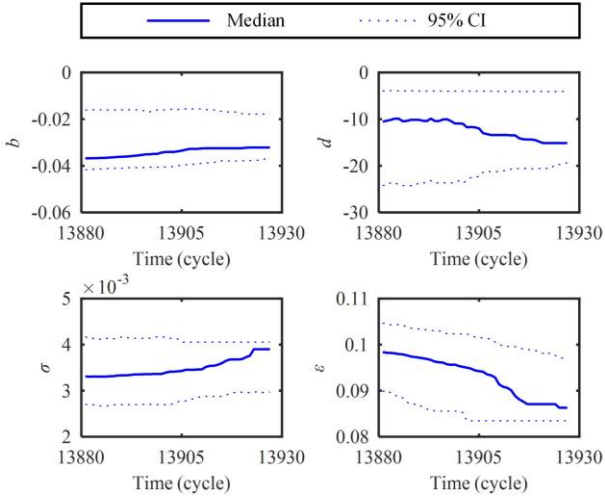
Fig. 7. Model parameter updating.
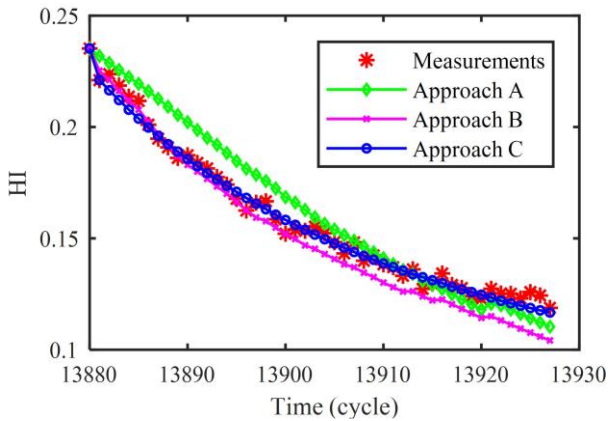
Fig. 8. Degradation tracking result comparison. (Approach A denotes the Paris model with standard PF, approach B denotes the exponential model with standard PF, and approach C denotes the proposed approach, i.e., improved double exponential model with HOPF)

Fig. 9 shows the absolute error of degradation tracking by different approaches. The average absolute errors of approach A, approach B, and the proposed approach are 0.0087, 0.0068,

and 0.0028, respectively. The HOPF-based approach gains a smaller tracking error than the standard PF. Those imply that the improved double exponential model can pick up the latent degradation trajectory better than the Paris model and the exponential model, and the HOPF can effectively filter the uncertainty of the measured HI.
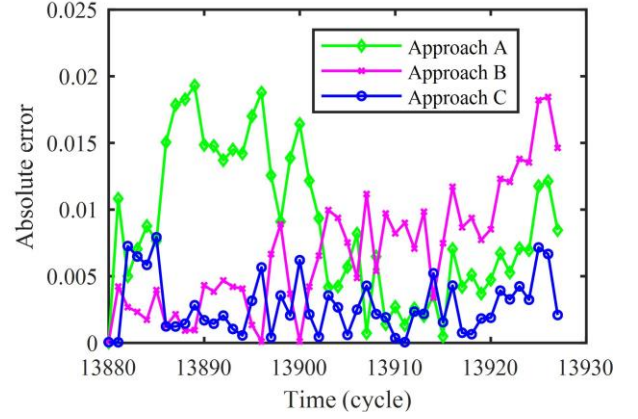
Fig. 9. Tracking error comparison.

The failure threshold is an essential part in RUL prediction of turbine disc. In this paper, the failure threshold is defined as the HI at the last cycle when the turbine disc is broken. The RUL prediction curve can be obtained by carrying out the RUL prediction task at each inspection time. Fig. 10 shows the RUL prediction curve. Here, the median is supposed as the RUL prediction value, and the 95% CI is chosen to evaluate the prediction uncertainty. We can see from Fig. 10 that the prediction value is close to the actual value and the 95% CI nearly falls into the 20% error interval of the actual RUL (range from 80% actual RUL to 120% actual RUL). The satisfied predictions demonstrate that the proposed HOPF-based approach performs well in RUL prediction for turbine disc.
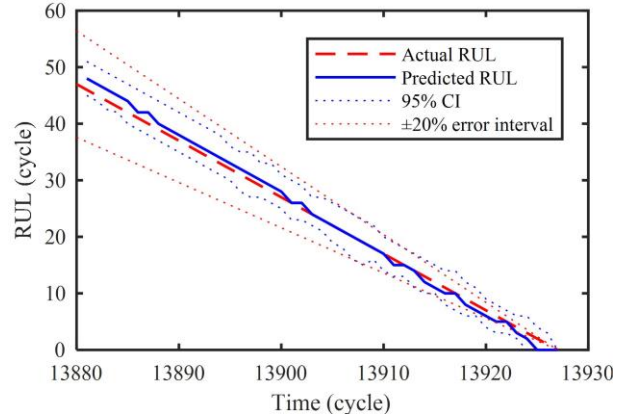
Fig. 10. RUL prediction at each inspection time.

Finally, another two PF-based approaches are also employed for comparison. As shown in Fig. 11, the PF-base approaches deviate from the actual RUL and has some volatility. The results imply that the improved double exponential model is more suitable than Paris model and exponential model for turbine disc degradation modeling. Two widely used assessment indexes, i.e., mean absolute error (MAE) and convergence (Con) [15], are employed here to quantitatively compare the prediction performance of different approaches. As shown in TABLE I, the proposed

approach gains the smallest MAE and Con values. That means the proposed perform the highest prediction accuracy and best prediction convergence among these approaches. The improvements of the proposed approach are mainly attributed to the high-order Markov assumption. The approach A and approach B consider that the degradation depends only on the previous state under the first-order Markov property. By contrast, the proposed approach develops an improved dual exponential model with the high-order Markov assumption, which consider the current degradation state depends not only on the previous state but also on multiple m-step-before states. Satisfied results demonstrate that high-order Markov assumption is more approximate for degradation modeling of turbine disc than the first-order assumption.
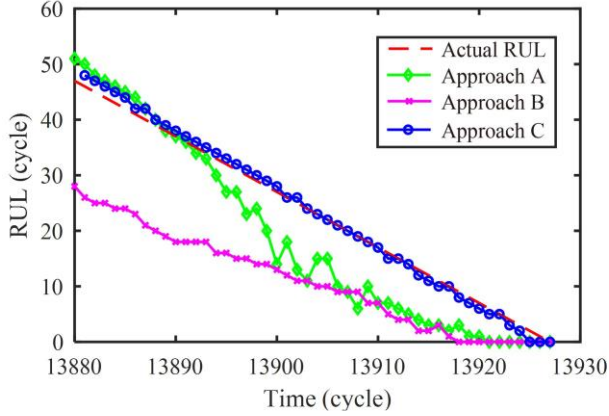
Fig. 11. RUL prediction comparison

TABLE I. QUANTITATIVE COMPARISON OF PREDICTION PERFORMANCE

Index	Approach A [13]	Approach B [14]	Proposed
MAE	5.7292	12.5417	0.8511
Con	24.1945	26.7489	21.9203

V. CONCLUSIONS

In order to overcome the two problems in RUL prediction of turbine disc, i.e., HI construction, and first order Markov limitation, this paper presents a HOPF-based RUL prediction approach for turbine disc. First, an HI is constructed for turbine disc based on the first shaft order vibration, which is a powerful feature of mass unbalance. Then, an improved double exponential model is developed to construct the HOSSM for describing the degradation process of turbine disc. Next, the parameters and state are updated with the help of HOPF. Finally, the RUL are predicted by a multi-step ahead prediction algorithm.

A low cycle fatigue test of turbine disc is provided to validate the proposed approach. The result shows that the established HI is capable to monitor the performance degradation process of turbine disc. Moreover, two PF-based approaches are employed to compare with the HOPF-based approach. The comparison results demonstrate that the HOPF-based approach performs better than PF-based approaches.

The subjective failure threshold seriously influences the RUL prediction accuracy. Few studies have been focused on adaptively setting failure threshold in RUL prediction. Therefore, adaptive threshold setting approach will be studied in our future research.

REFERENCES

[1]   Z. Lv, H. Huang, H. Gao, F. Zuo, and H. Wang, "Lifetime prediction for turbine discs based on a modified Walker strain model," Journal of Mechanical Science and Technology, vol. 29, pp. 4143-4152, 2015.

[2]   S. Zhu, Y. Liu, Q. Liu, and Z. Yu, "Strain energy gradient-based LCF life prediction of turbine discs using critical distance concept," International Journal of Fatigue, vol. 113, pp. 33-42, 2018.

[3]   V. Shlyannikov, and I. Ishtyryakov, "Crack growth rate and lifetime prediction for aviation gas turbine engine compressor disk based on nonlinear fracture mechanics parameters," Theoretical and Applied Fracture Mechanics, vol. 103, pp. 1-14, 2019.

[4]   Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," Mechanical Systems and Signal Processing, vol. 104, pp. 799-834, 2018.

[5]   F. Ahmadzadeh, and J. Lundberg, "Remaining useful life estimation: review", International Journal of System Assurance Engineering and Management, vol. 5, pp. 461-474, 2014.

[6]   K. Tsui, N. Chen, Q. Zhou, Y. Hai, and W. Wang, "Prognostics and health management: A review on data driven approaches", Mathematical Problems in Engineering, vol. 2015, pp. 1-17, 2015.

[7]   H. Hanachi, C. Mechefske, J. Liu, A. Banerjee, and Y. Chen, "Performance-based gas turbine health monitoring, diagnostics, and prognostics: A survey," IEEE Transactions on Reliability, vol. 67, pp. 1340-1363, 2018.

[8]   K. Peng, R. Jiao, J. Dong, and Y. Pi, "A deep belief network based health indicator construction and remaining useful life prediction using improved particle filter," Neurocomputing, vol. 361, pp. 19-28, 2019.

[9]   C. Yang, Q. Lou, J. Liu, Y. Yang, and Y. Bai, "Particle filter-based method for prognostics with application to auxiliary power unit," Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, pp. 198-207, 2014.

[10] W. Wang, "Disk crack detection and diagnosis for gas turbine engines", 2006 IEEE Aerospace Conference, 2006, pp. 1-12.

[11] Y. Chang, and H. Fang, "A hybrid prognostic method for system degradation based on particle filter and relevance vector machine," Reliability Engineering & System Safety, vol. 186, pp. 51-63, 2019.

[12] C. Matteo, S. Claudio, M. Andrea, and G. Marco, "Real-time prognosis of crack growth evolution using sequential Monte Carlo methods and statistical model parameters", IEEE Transactions on Reliability, vol. 64, pp. 736-753, 2015.

[13] N. Li, Y. Lei, Z. Liu, and J. Lin, "A particle filtering-based approach for remaining useful life prediction of rolling element bearings", 2014 IEEE Conference on Prognositcs and Health Management, 2014.

[14] D. An, J. Choi, and N. Kim, "Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab", Reliability Engineering & System Safety, vol. 115, pp. 161-169, 2013.

[15] S. Abhinax, C. Jose, S. Bhaskar, S. Sankalita, and G. Kai, "Metrics for offline evaluation of prognostic performance", International Journal of Prognostics and Health Management, vol. 1, pp. 4-23, 2010.

Sub-Sampled Two-Dimensional SAR Imaging Method Based on MIMO FMCW Radar

Zesheng Ren
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
Silence-SCIENCE@sjtu.edu.cn

Yuyong Xiong
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
yy.xiong@sjtu.edu.cn

Songxu Li
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
songxu.li@sjtu.edu.cn

Dong Wang
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
dongwang4-c@sjtu.edu.cn

Zhike Peng *
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
z.peng@sjtu.edu.cn

Abstract—**One challenge of near field millimeter wave imaging technology is to solve the conflict between high image resolution and high computational cost with raw data. This paper investigates the use of MIMO FMCW radar in two-dimensional synthetic aperture radar imaging process under sparsely sampling conditions at space axes. Exploiting the thought that considering the imaging process as a filter, multiple virtual channels' information is taken advantaged with complex gains so that even sub-sampled data blocks can successfully rebuild the target image with little aliasing components. Simulation validations are provided to show the effectiveness of this method.**

Keywords—FMCW, MIMO radar, subsampling, SAR imaging

I. INTRODUCTION

Microwave radars are commonly used for target detection and motion measurement [1]-[2]. And the imaging technology with millimeter wave (mm Wave) radar has played an important role in many applications including like, structure inspection [3], weapon detection [4], medical monitoring [5]. As one of these imaging methods, raw data of synthetic aperture radar (SAR) sampled both in space and time axes must be properly processed before the final image turns out.

Conducting sub-sample on a time series signal, i.e., the sampling frequency is lower than two times of the maximum frequency components of the original signal, will lead to aliasing phenomenon in frequency domain. Similarly, synthetizing the aperture is also a kind of sampling process, not only at time axis but also at space axes. It also follows the Nyquist sampling theorem that spatial sampling rate must be two times bigger than the band limits of target. In previous two-dimensional SAR imaging methods, signal aliasing in space domain because of the sub-sampling will lead to extra image components called ghost image [6]. The ghost images caused by aliasing components will duplicate around the real target image after nonlinear operation on power strength and shift distance.

On the other hand, sampling rate that satisfies the Nyquist theorem in two-dimensional SAR imaging method will result in big data blocks, two dimensions in space domain and one dimension in time domain, which means large amounts of computation. In order to solve the conflict between need of high-resolution SAR images and data process complexity, we propose a different SAR imaging method based on MIMO FMCW radar. In this method, we take advantage of information from different virtual channels of MIMO radar to reinforce the image result so that even under a sparsely sampling rate in space domain, ghost images by aliasing components can be successfully eliminated.

To use the information from different virtual channels, we assume the imaging process as a linear filter , 3D data blocks from virtual channels as input and the target image required as output. We add proper complex gain to signal blocks from virtual channels before going through the filter above so that the output will not contain duplication of real target image. Simulation results also show that this sparsely sampled MIMO SAR imaging method do works well on erasing the ghost images.

The rest of this paper is organized as follows. In section II, a brief introduction of the subsampled MIMO SAR imaging principle is given. Section III presents the simulation and analysis results. A summary of the contribution of this paper is provided in Section IV.

II. SUBSAMPLED MIMO SAR IMAGING METHOD

For a typical kind of FMCW radar transmitting signal, single transmit antenna signal can be express in complex form as:

$$m(t) = e^{j2\pi(f_0 t + 0.5 K_r t^2)} \tag{1}$$

Where f_0 is the carrier frequency, K_r is the frequency module rate. The baseband signal of a single scatter at distance R from antenna can be written as:

$$s(t) = \sigma e^{j2\pi(f_0\tau + K_r\tau t - 0.5 K_r \tau^2)} \tag{2}$$

Where τ is the time delay, sigma express the microwave power strength received by the radar:

$$\sigma = p/R^2 \tag{3}$$

p is the target reflectivity, timed by $1/R^2$ representing the distance decay [7]. The last phase term in (2) is known as the residual video phase (RVP) and can be neglected in the imaging process. Rewrite formula (2) into wave number form, i.e., let $k = 2\pi f/c$, we can derive the beat frequency form:

$$s(k) = p\frac{e^{j2kR}}{R^2}, \quad \frac{2\pi f_0}{c} \le k \le \frac{2\pi f_T}{c} \tag{4}$$

Where f_T is the end frequency. In this formula, distance R between radar channels and target is a function of space coordinates x and y at the plane radar's synthetic aperture, which means the beat signal received in 2D SAR imaging process is a function of x and y, and wave number k. Apply (4) on every target point at target plane (parallel with the synthetic aperture) or their rejections on the plane, according to the linear system assumption, we can get the final beat signal block form as below:

$$s(x, y, k) = \iint p(x', y') \frac{e^{j2kR}}{R^2} dx' dy' \qquad (5)$$

x' and y' are spatial coordinates at target plane and reflectivity p is a 2D function defined on it. Then the total beat signal from the target plane is expressed as a combination of reflectivity p, phase factor e^{j2kR} and distance R from transceiver element to the target. Obviously, the image of the target plane we want to obtain is indeed the 2D reflectivity function on target plane, $p(x, y)$. And the purpose of our SAR imaging method is to rebuild $p(x, y)$ from the beat signal data block according to the equation (5).

All the process above is used in SISO radar system. When it comes to MIMO radar situation, we treat the imaging method in SISO as a signal filtering process. The input of the filter is the 3D data block of beat signal and the output of the filter is the target image we required. This imaging process can be written as a simple multiplication form in frequency domain as:

$$P(kx, ky) = \int S(kx, ky, k) \cdot H dk \qquad (6)$$

Where H is the filter we built from SISO radar SAR imaging method. Just as mentioned at the beginning, usually enough spatial sample rate need to be guaranteed to avoid aliasing components and the ghost image that follows. To erase the ghost image, the beat signal blocks from MIMO radar's different virtual elements will be taken advantage and the rewritten imaging form in equation (5) makes it easy to derive the new form of imaging when using multiple virtual channels as:

$$P(kx, ky) = \int \sum [w_i \, S_i(kx, ky, k)] H dk \qquad (7)$$

Where w_i is the complex gain added to the ith virtual channel's data block before go through the filter. By choosing proper complex gain and enough virtual channels, relative simulation results show that it successfully reduces the ghost images effect in final target image $p(x, y)$.

A reasonable thought about this imaging method is that the more virtual channels are involved in the imaging process, the better the results of final target image will be. However, simulation results show that if too much virtual channels is involved, the image result can be as worse as traditional imaging method under sub-sampling condition in space.

III. VALIDATION AND EVALUATION

The sub-sampled 2D SAR imaging method with MIMO radar in this paper have been simulated with MATLAB R2019a software. As a prove of the imaging method, we use four scatters as target and apply traditional 2D SAR-imaging method with proper spatial sampling rate that satisfies Nyquist theorem firstly. Then we simulated the sparsely sampling situation using traditional method and our new

method with MIMO radar respectively. Simulation results turn out that this sub-sampled MIMO SAR method successfully reduce the ghost image components from real target image result.

All the simulation results in this paper shares the same radar physical parameters and FMCW parameters.

We choose carrier frequency at 77.38GHz with modulating slope at 6.337×10^{13}Hz/s for FMCW waves in our simulation. The synthetic aperture size is 200mm at both x and y axes. And the target plane is settled at 300mm away from the aperture plane.

A. Proper Spatial Sampled Data with Traditional 2D SAR Simulation

We set the radar moving step at 0.8mm on both axes in space so that satisfies Nyquist theorem. Fig.1 presents the simulation result that shows four target scatters clearly in the final image and scatters' position is clearly demonstrated
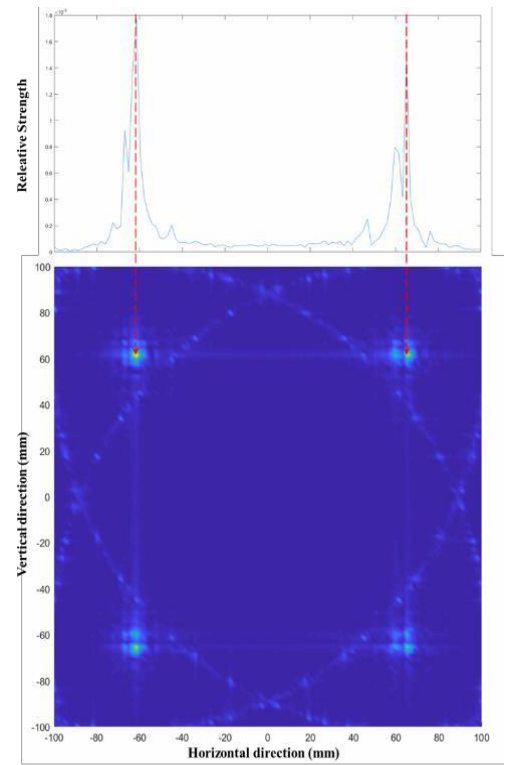
Fig. 1. Proper sampled scatters' image on targe plane

B. Sparsely Spatial Sampled Data with Traditional 2D SAR Simulation

The radar moving step is set at 8mm on both axes in space so that far away from required Nyquist sampling rate. And the simulation result shows that the target scatters' image totally submerged into the ghost images by aliasing components. The results can hardly tell the target points' original position.

Fig.2 shows that the whole image is covered by meaningless ghost images induced by aliaing componnets that totally cover up the real target.

C. Sparsely Spatial Sampled Data with MIMO SAR Simulation

Similar to the scenario in B, moving step is set at 8mm on both axes in space but three virtual channels are used in the imaging process with proper complex gains. Simulation shows that the target scatters' image again clearly appears and

the ghost image has been greatly reduced so that scatters can be located precisely. A difference from the proper sampled result is that both the main and minor lobe become little wider and shifted slightly thus spatial resolution and accuracy is influenced accordingly.
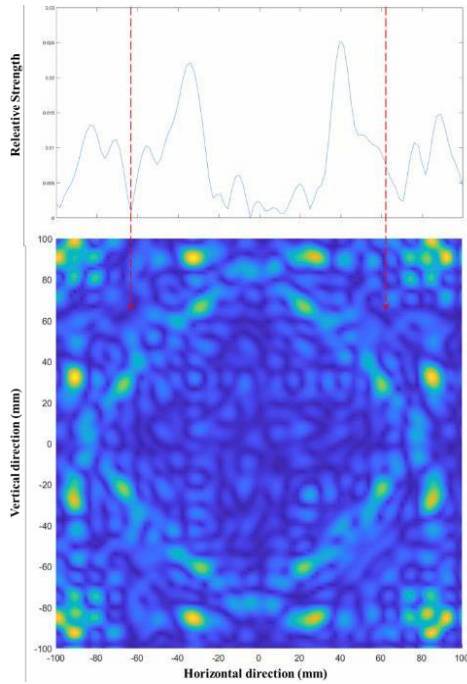
Fig. 2. Subsampled scatters' image on target plane
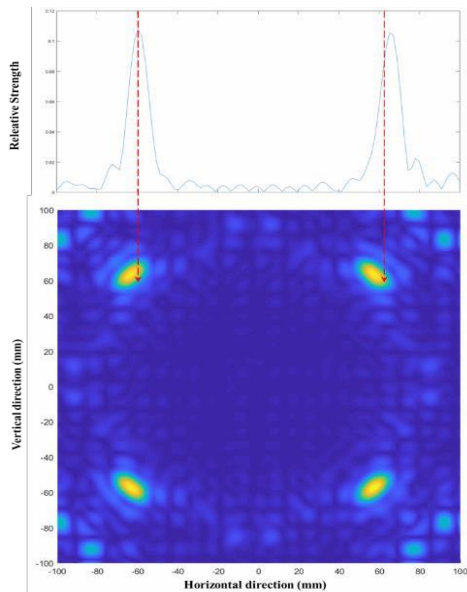
Fig. 3. Subsampled scatters' image by MIMO SAR method

Fig.3 shows that the subsampled scatters' image by MIMO SAR method which real target appears again with much more inconspicuous ghost image compared with the situation in Fig.2. Besides, relevant simulation results show that imaging results also depend on the positions of virtual channels involved in the process and the number of the virtual channels can neither too less nor too much, otherwise there still can be ghost images left or even a loss of real target.

IV. Conclusion

In this paper the capability of sub-sampled MIMO-SAR imaging method is demonstrated that it can effectively reduce the ghost images caused by aliasing components. Based on the multiple information from different virtual channels with proper complex gains and the input-filter-output thought about the imaging process, our approach method performs well under sparsely sampling conditions in space with little compromise of lobe width and resolution. Simulation result of scatter target points validate the effectiveness on cancelling duplicating components from final image.

References

[1] Y. Xiong, Z. Peng, G. Xing, W. Zhang and G. Meng, "Accurate and Robust Displacement Measurement for FMCW Radar Vibration Monitoring," IEEE Sensors Journal, vol. 18, no. 3, pp. 1131-1139, Feb. 2018.

[2] Y. Xiong, S. Chen, X. Dong, Z. Peng and W. Zhang, "Accurate Measurement in Doppler Radar Vital Sign Detection Based on Parameterized Demodulation," IEEE Transactions on Microwave Theory and Techniques, vol. 65, no. 11, pp. 4483-4492, Nov. 2017.

[3] S. Kharkovsky, J. T. Case, M. A. Abou-Khousa, R. Zoughi, and F. L. Hepburn, ''Millimeter-wave detection of localized anomalies in the space shuttle external fuel tank insulating foam,'' IEEE Trans. Instrum. Meas., vol. 55, no. 4, pp. 1250–1257, Aug. 2006.

[4] D. M. Sheen, D. L. McMakin, and T. E. Hall, ''Three-dimensional millimeter-wave imaging for concealed weapon detection,'' IEEE Trans. Microw. Theory Techn., vol. 49, no. 9, pp. 1581–1592, Sep. 2001.

[5] L. Chao, M. N. Afsar, and K. A. Korolev, ''Millimeter wave dielectric spectroscopy and breast cancer imaging,''in Proc. 7th Eur. Microw. Integr. Circuit Conf., Amsterdam, The Netherlands, Oct. 2012, pp. 572–575.

[6] M. Yanik, and M. Torlak, "Near-Field MIMO-SAR Millimeter-Wave Imaging with Sparsely Sampled Aperture Data," IEEE Access, vol. 7, pp. 31801–31819, Mar. 2019.

[7] A. Meta, P. Hoogeboom, and L. P. Ligthart, ''Signal processing for FMCW SAR,'' IEEE Trans. Geosci. Remote Sens., vol.45, no.11, pp.3519–3532, Nov. 2007.

Principal Component Analysis Based Kullback-Leibler Divergence for Die Cracks Detection

Sha Wei
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
weisha@sjtu.edu.cn

Dong Wang
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
dongwang4-c@sjtu.edu.cn

Zhike Peng*
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
z.peng@sjtu.edu.cn

Abstract—Die cracks are a vital issue that directly influences the quality of chip assemblies. In this paper, we focus on detecting die cracks using principal component analysis (PCA) and Kullback-Leibler (K-L) divergence. Our method involves data fusion, including three steps: 1) apply PCA to convert high-dimensional data to low-dimensional data; 2) obtain the frequency distribution histograms of the transformed data and fit them; 3) use K-L Divergence based state index to quantitatively evaluate die cracks. Our method works very well with real-life data. Die cracks are identified according to die cracks data showing skewed distribution, while normal data have Gaussian distribution. Moreover, the proposed state index could successfully detect die cracks.

Keywords—PCA, data fusion, Gaussian distribution, die cracks, K-L divergence, state index

I. INTRODUCTION

Semiconductor chips are widely used in various fields and have occupied a critical position of economic development. In recent years, the rapid growth in portable electronic products has challenged the chips with high reliability [1,2]. The quality of chips largely depends on the quality of die. Die processing procedure involves many processes, which include surface mount technology, die prepare, die attach, wire bond, etc. Herein, we concentrate on die prepare process. The purpose of the die prepare process is to convert wafer level to die level. This process primarily consists of back grinding and wafer saw. The former aims to reduce the die thickness using a grinding wheel and the latter transforms wafers into hundreds or thousands of single dies. Die cracks are sometimes caused by these mechanical procedures [3]. The cracks may extend to the metal layer and affect the quality of the die, which greatly reduces the output of products, resulting in economic losses.

There are many technologies for die cracks detection which mainly include contact and non-contact. Liu et al. proposed the impact test technology as a representative of the contact method [4]. This method could detect die cracks through the impact test, which requires simple test equipment and have short time consuming. However, this method need contact semiconductor chips that may cause potential damage. Thus, many non-contact detection methods are widely used, which mainly contain eddy current sensor technology, terahertz imaging technology, scanning acoustic microscopy technology, light scattering technology, line-laser phase-locked thermal imaging technology. Eddy current sensor technology [5] uses the principle of electromagnetic induction between the electromagnetic field and the metal to detect, which means this method is only used for conductive materials. Besides, strong eddy current will have adverse effects on chip performance. Terahertz imaging technology utilizes terahertz wave of special band to detect die cracks [6]. This method has a relatively high temporal and spatial resolution, which could detect the microcrack. However, this method has a shallow penetration depth and cannot penetrate the metal layer. Scanning acoustic microscopy technology detects die cracks based on the principle of different degrees of acoustic absorption and reflection at die cracks [7,8]. This method penetrates deeper than terahertz imaging technology, but the detection time is longer, which makes it unsuitable for online detection. The principle of light scattering technology is based on the annular illumination and the Fourier transform optical system [9]. It shows that micro-scratches with the depth of only 10nm could be detected independent of their directions. However, this method is subject to the die surface and light conditions. Line-laser phase-locked thermal imaging technology does not rely on historical data and is used to instantaneously detect die cracks, but is affected by the direction of line scanning [10].

It is expensive to detect die cracks from hardware. Thus, the purpose of this study is to find an effective method to detect die cracks from analyzing signals collected during chip production. Firstly, we present an overview of methodology proposed. Then, a practical example is shown to prove the effectiveness of the proposed method. Finally, we conclude some criteria to detect die cracks.

II. THEORETICAL BASIS

To reduce economic losses, we analyze the collected signals. Generally, these types of signals range widely and the degree of crack is weak, which makes defects hard to detect. Taking data fusion into account, we employ principal component analysis (PCA) to deal with these multichannel signals. Then, the histograms of frequency distribution of principal components are drawn and fitted. At last, we introduce Kullback-Leibler divergence and state index to detect die cracks quantificationally.

A. Theory of PCA

The PCA is a statistical procedure that uses an orthogonal transformation to convert a set of strong correlated variables into a set of uncorrelated or weak correlated variables which contain as much information about the original variables as possible [11,12]. The variable after transformation is called principal components. Therefore, PCA, a simple and convenient method, has the following features: i) simplify the data by dimensionality reduction; ii) save as much as possible of the original signal information [13]. By dimension reduction, the number of variables to be analyzed could be reduced and the coupling between variables could be eliminated or weakened. The PCA specific algorithm steps are as follows

---

*Zhike Peng is the corresponding author. (e-mail: z.peng@sjtu.edu.cn).

Algorithm 1

1. Standardized the data;
2. Calculate the covariance matrix of the standardized data;
3. Calculate the singular values and eigenvectors of the covariance matrix;
4. Keep the most important k features;
5. Operate inverse transform to obtain dimensionality reduction variables;
6. Obtain the frequency distribution histogram and fit probability density function of dimensionality reduction variables;
7. Calculate K-L Divergence based state index.
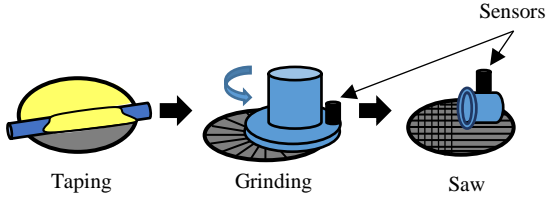
Fig. 1. The schematic diagram of the experiment.

An input signal of n components is observed, as expressed in (1).

$$\mathbf{X} = \left(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n \right)^{\mathrm{T}} \qquad (1)$$

Firstly, estimate the mean value of $\mathbf{x}_i$, denoted as $\overline{\mathbf{x}}_i$. Then calculate the standard deviation ζ. Thus, the standardization matrix can be written as

$$\overline{\mathbf{X}} = \left(\frac{\mathbf{x}_1 - \overline{\mathbf{x}}_1}{\zeta}, \frac{\mathbf{x}_2 - \overline{\mathbf{x}}_2}{\zeta}, ..., \frac{\mathbf{x}_n - \overline{\mathbf{x}}_n}{\zeta} \right)^{\mathrm{T}} \qquad (2)$$

PCA optimization problems are usually measured by the sum of squared deviations or variances. The solution methods include singular value decomposition (SVD) and non-linear iterative partial least squares (NAIPALS) [14]. Herein, we have chosen SVD. For SVD, we calculate the covariance matrix $\mathbf{S}$ of $\overline{\mathbf{X}}$. Then, take the SVD for the covariance matrix $\mathbf{S}$. It can be expressed as

$$\mathbf{S} = \mathbf{U}\Sigma\mathbf{V}^* \qquad (3)$$

where $\mathbf{U}$ is a $n \times n$ unitary matrix, Σ is a $n \times k$ diagonal matrix, which consists of singular values, $\mathbf{V}$ is a $n \times k$ unitary matrix, superscript * stands for conjugate transpose. Extract the first m singular values, and the first m columns of $\mathbf{U}$ can be spanned into an Eigen-subspace. Lastly, take the inverse operation to reconstruct the 1-D signal $\mathbf{Z}$ with the same size of the original signal. The dimensionality reduction matrix can be obtained by

$$\mathbf{Z} = \left(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_d \right)^{\mathrm{T}} \qquad (4)$$

where $\mathbf{z}_1$ is the first principal component, $\mathbf{z}_2$ is the second principal component.

B. Theory of Gaussian Distribution

If the random variable $\mathbf{X}$ obeys a Gaussian distribution with a mathematical expectation of μ and a variance of σ^2,
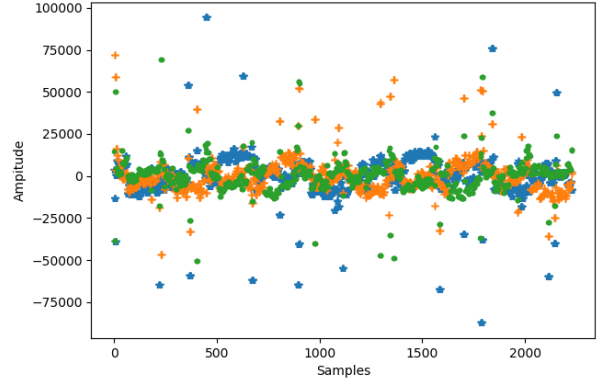
Fig. 2. The first three principal components of normal data (the blue *, orange +, and green . stand for the first principal component, the second principal component and the third principal component, respectively).
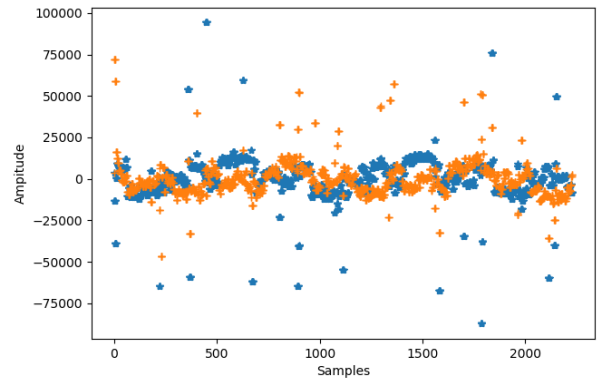
Fig. 3. The first two principal components of normal data (the blue * and orange + represent the first principal component and the second principal component, respectively).

it is denoted as $N\left(\mu, \sigma^2\right)$. The probability density function is defined as [15]

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right) \qquad (5)$$

Standard Gaussian distribution is a special case when $\mu = 0$ and $\sigma = 1$, as expressed in (6)

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} \right) \qquad (6)$$

Actually, μ is the position parameter, which describes the central tendency position of Gaussian distribution. The Gaussian distribution takes $x = \mu$ as its axis of symmetry. σ is the shape parameter, which describes the degree of dispersion of Gaussian distribution [16]. The larger the σ, the more dispersed the data and the flatter the curve, vice versa.

C. Theory of Kullback-Leibler Divergence

Kullback-Leibler divergence is a measure of the asymmetry of the difference between two probability distributions [17]. Let $P(x)$ and $Q(x)$ be two probability distributions on the random variable $\mathbf{X}$, then in the case of discrete random variables, the K-L divergence definitions is
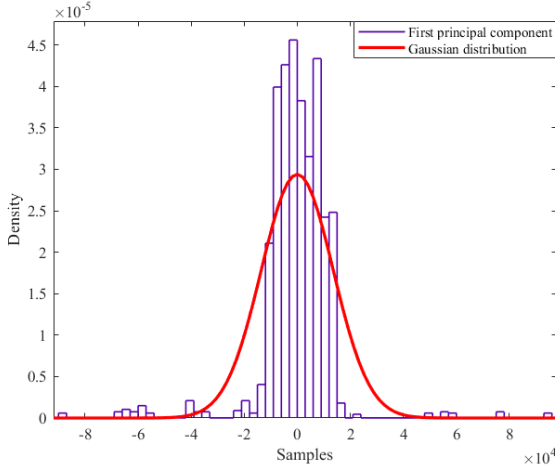
Fig. 4. The frequency distribution histogram and fitted probability density fuction of the first principal component of normal data.
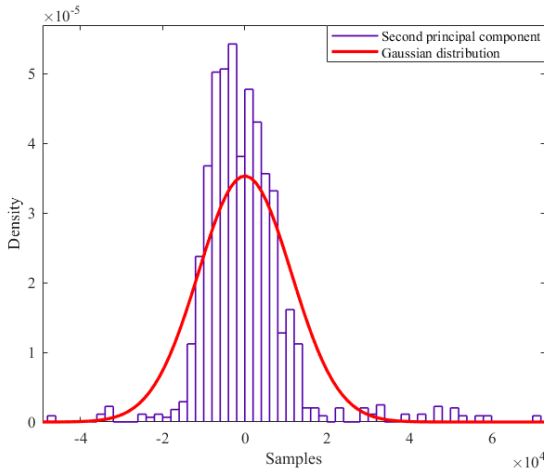
Fig. 5. The frequency distribution histogram and fitted probability density fuction of the second principal component of normal data.

$$\mathrm{KL}\left(P \| Q\right) = \sum P(x) \log \frac{P(x)}{Q(x)} \qquad (7)$$

K-L divergence could measure the distance between two random distributions. When two random distributions are the same, their K-L divergence is almost zero . Otherwise, the greater the difference between two random distributions, the greater their K-L divergence.

In order to quantitatively evaluate K-L divergence, a state index (SI) is proposed. The definition of this SI is as follows.

$$\mathrm{SI} = \frac{\mathrm{KL}\left(C \| A\right)}{\mathrm{KL}\left(C \| A\right) + \mathrm{KL}\left(C \| B\right)} \qquad (8)$$

where A, B and C represent Gaussian distribution of normal data, abnormal data and test data, respectively.

The complete algorithm is outline in Algorithm 1.

III. DEMONSTRATION

In the process of back grinding and wafer saw, a variety of sensors are mounted on shafts to collect signals. The
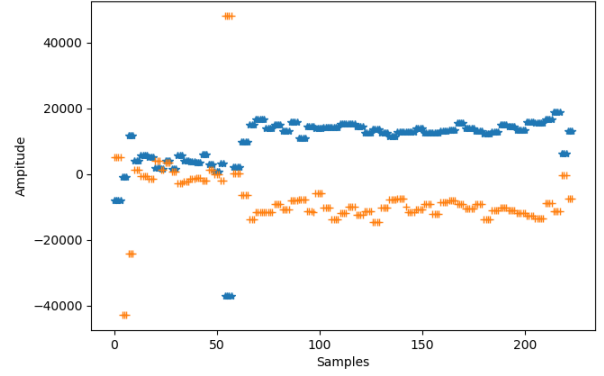
Fig. 6. The first two principal components of crack data (the blue * and orange + represent the first principal component and the second principal component, respectively).

schematic diagram of the experiment is shown in Fig. 1. This section is includes two parts: one is separated data, the other is mixed data.

A. Separated Data

In this case, normal data and crack data are separated. We select 16 channels of data related to rpm, torque, and wear as the input of the PCA. PCA is performed on the normal data first, and the crack data are processed using the obtained coefficients. We notice that the first three principal components are quite similar, as illustrated in Fig. 2. To reduce redundancy and without lose generality, we finally choose two principal components to express the results after dimensionality reduction. The first and second principal components of normal data are illustrated in Fig. 3. The principal components of normal data are fluctuant, as shown in Fig. 3. Then, we obtain the frequency distribution histograms of the first and second principal components and fit their probability density function, as illustrated in Fig. 4 and Fig. 5, respectively. Fig. 4 and Fig. 5 show that the principal components of normal data belong to Gaussian distributions, because the values of log likelihood are -24384.3 and -23974.5, respectively.

TABLE I. THE VALUES OF LOG LIKELIHOOD

The Values of Log Likelihood	Fitted Probability Distribution	
	Gaussian distribution	*t Location-Scale distribution*
The first principal component	-2340.09	-2245.15
The second principal component	-2382.05	-2287.61

TABLE II. K-L DIVERGENCE AND THE STATE INDEX OF TEST DATA FROM NORMAL DATA

Test Data from Normal Data	K-L Divergence and the State Index		
	KL(test‖normal)	*KL(test‖crack)*	*The state index*
The first group	2.668017682	25.81028560	0.09368598
The second group	9.159709586	20.67863216	0.30697784
The third group	5.612490280	17.70611863	0.24068718
The fourth group	1.724643101	24.06646447	0.06686968
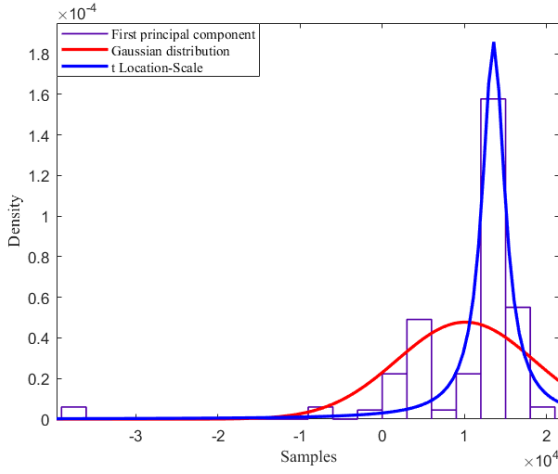
Fig. 7. The frequency distribution histogram and fitted probability density fuction of the first principal component of crack data.
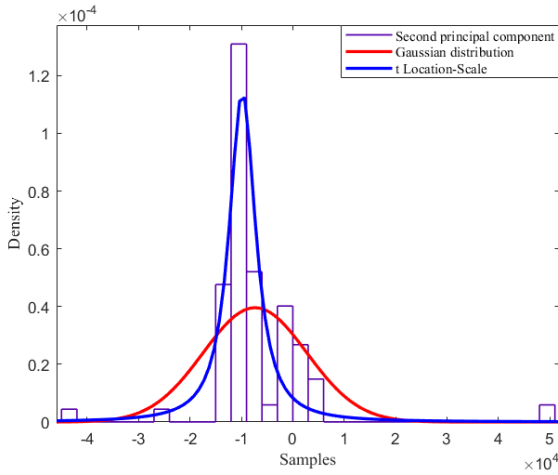
Fig. 8. The frequency distribution histogram and fitted probability density fuction of the second principal component of crack data.

The first and second principal components of crack data are illustrated in Fig. 6. As shown in Fig. 6, the principal components of crack data are almost constant. Then, we obtain the frequency distribution histograms and fitted curves of the first principal component and the second principal component, as illustrated in Fig. 7 and Fig. 8, respectively. We use t Location-Scale distribution and Gaussian distribution to fit the principal components, as shown in the Fig. 7 and Fig. 8. The values of log likelihood are listed in Table I. The Gaussian distribution is close to the t Location-Scale distribution for the values of log likelihood. Thus, we select the Gaussian distribution to fit the histogram. The fitted curve of the first principal component is longer to the left and shorter to the right, which is called left skewed distribution. The fitted curve of the second principal component is longer to the right and shorter to the left, which is called right skewed distribution.

TABLE III. K-L DIVERGENCE AND THE STATE INDEX OF TEST DATA FROM CRACK DATA

Test Data from Crack Data	K-L Divergence and the State Index		
	KL(test‖normal)	KL(test‖crack)	The state index
The first group	40.56476913	7.082519966	0.85135524
The second group	63.22512898	4.127164163	0.93872274
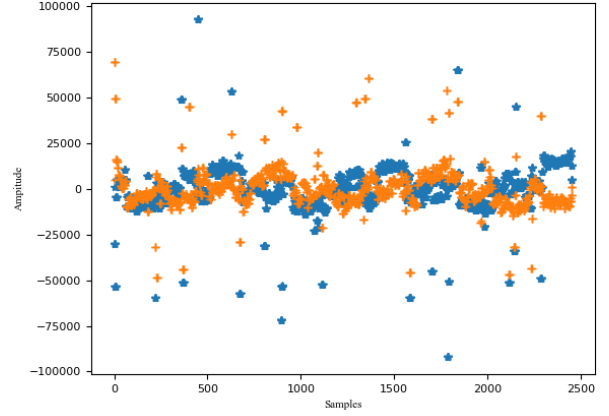
Fig. 9. The first two principal components of mixed data (the blue * and orange + represent the first principal component and the second principal component, respectively).

On the basis of establishing the distribution models of normal data and crack data, test data are introduced to identify die cracks or not by calculating the state index. For simplicity, only the first principal component of normal data and crack data are analyzed. Considering the difference in the number of samples, the test data from the first principal component of normal data are a group of every 500 sampling points, and there is a total of 4 groups of test data, as illustrated in Table II. The test data from the first principal component of crack data are a set of every 150 sampling points, with the overlap rate of 0.5, and there are two groups in total, as illustrated in Table III. According to Table II and Table III, we could draw the conclusion that when the test data are similar to normal data, the state index is small, vice versa. After actual testing, we choose 0.5 as the threshold of state index to detect die cracks.

B. Mixed Data

In this case, normal data and crack data are mixed. We also select the same 16 relatively critical channel signals as the input to PCA. PCA is performed on mixed data. As mentioned above, we finally choose two principal components to express the results. The first principal component and the second principal component of mixed data are illustrated in Fig 9. We could still clearly detect die cracks from Fig 9. It is noted that the crack condition cannot be recognized from the fitted curve of mixed data due to normal data and crack data mixed together.

IV. CONCLUSIONS

In this paper, an algorithm based on PCA and Kullback-Leibler divergence is presented to successfully detect die cracks. The principal components of dimensionality reduction normal data are fluctuant while almost constant of dimensionality reduction crack data. Simultaneously, the distribution of dimensionality reduction normal data is Gaussian distribution while the distribution of dimensionality reduction crack data is skewed distribution. We could determine the type of test data by comparing the value of K-L Divergence based state index with 0.5, which means that when the value of state index is less than 0.5, the test data are normal data, and vice versa. In practical industrial applications, we can detect crack data according to these three criteria. It is also possible to design statistical process control based on these results to detect heterogeneous states more scientifically.

REFERENCES

[1] W. Kroeninger, L. Schneider, and G. Wagner, "Creating Stable and Flexible Chips for Thin Packages," Semiconductor International, vol. 27, no. 12, pp. p.59-60,62,64, 2004.

[2] N. Mclellan, N. Fan, S. Liu, K. Lau, and J. Wu, "Effects of Wafer Thinning Condition on the Roughness, Morphology and Fracture Strength of Silicon Die," Journal of Electronic Packaging, vol. 126, no. 1, pp. 110-114, 2004.

[3] D. Y. R. Chong, W. E. Lee, B. K. Lim, J. H. L. Pang, and T. H. Low, "Mechanical failure strength characterization of silicon dice," in Electronics Packaging Technology, Conference, 2003.

[4] J. Liu, "Research Progress of Fatigue Failure Prediction Methods and Damage Mechanism," Journal of Mechanical Engineering, vol. 50, no. 20, p. 26, 2014.

[5] D. J. Sadler and H. A. Chong, "On-chip eddy current sensor for proximity sensing and crack detection," Sensors & Actuators A, vol. 91, no. 3, pp. 340-345, 2001.

[6] quot et al., "Terahertz Spectroscopy and Imaging," Infrared & Laser Engineering, vol. 171, no. 2, pp. 359-368, 2013.

[7] P. Dong, "On-chip ultra-fast data acquisition system for optical scanning acoustic microscopy using 0.35um CMOS technology," University of Nottingham, 2009.

[8] A. Belyaev, O. Polupan, S. Ostapenko, D. Hess, and J. P. Kalejs, "Resonance ultrasonic vibration diagnostics of elastic stress in full-size silicon wafers," Semiconductorence & Technology, vol. 21, no. 3, pp. 254-260, 2006.

[9] S. Takahashi, T. Miyoshi, Y. Takaya, and T. Abe, "New optical measurement technique for Si wafer surface defects using annular illumination with crossed nicols," in ASPE Proc, 2002.

[10] Y. K. An, M. K. Ji, and H. Sohn, "Laser lock-in thermography for detection of surface-breaking fatigue cracks on uncoated steel structures," Ndt & E International, vol. 65, no. jul., pp. 54-63, 2014.

[11] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," Chemometrics and intelligent laboratory systems, vol. 2, no. 1-3, pp. 37-52, 1987.

[12] H. Abdi and L. J. Williams, "Principal component analysis," Wiley interdisciplinary reviews: computational statistics, vol. 2, no. 4, pp. 433-459, 2010.

[13] K. J. Deluzio, U. P. Wyss, B. Zee, P. A. Costigan, and C. Serbie, "Principal component models of knee kinematics and kinetics: Normal vs. pathological gait patterns," Human Movement Science, vol. 16, no. 2-3, pp. 201-217, 1997.

[14] D. Michie, D. J. Spiegelhalter, and C. Taylor, "Machine learning," Neural and Statistical Classification, vol. 13, no. 1994, pp. 1-298, 1994.

[15] A. Lyon, "Why are normal distributions normal?," The British Journal for the Philosophy of Science, vol. 65, no. 3, pp. 621-649, 2014.

[16] D. G. Altman and J. M. Bland, "Statistics notes: the normal distribution," Bmj, vol. 310, no. 6975, p. 298, 1995.

[17] S. Kullback and R. A. Leibler, "On information and sufficiency," The annals of mathematical statistics, vol. 22, no. 1, pp. 79-86, 1951.

Research on Inductively Coupled Full Duplex Communication Method with Power Transmission

Tao Zhang
College of Mechanical Engineering
Chongqing University
Chongqing, China
1160412805@qq.com

Deqi Zhang
College of Mechanical Engineering
Chongqing University
Chongqing, China
1718265911@qq.com

Wenbin Huang*
College of Mechanical Engineering
Chongqing University
Chongqing, China
whuang@cqu.edu.cn

Abstract—**In some occasions, Inductively Coupled Power Transfer (ICPT) systems are required to transmit bidirectional data. In order to realize this purpose, a full duplex data transmission method is proposed in this paper. The power transmission system adopts the inductor/capacitor/capacitor-series (LCC-S) compensation structure to realize the constant voltage output of secondary coil. In the signal transmission system, the forward and reverse signal transmission systems adopt Amplitude Modulation (AM) technology and Amplitude Shift Keying (ASK) technology, respectively. And a frequency selective network is added to avoid interference between forward and reverse signal transmission. Then the feasibility of the scheme is verified theoretically by circuit simulation software Multisim, and the optimal element parameters are selected according to the calculation and simulation results. Finally, a hardware circuit is built to thoroughly verify the feasibility of the scheme.**

Keywords—inductively coupled power transfer (ICPT), amplitude modulation of electric energy, high frequency injection signal transmission, full duplex communication

I. Introduction

Inductive coupling power transmission has advantages of simple and flexible design, stable and reliable power transmission, and excellent anti-interference ability. It has a great application prospect in implantable medical equipment, wireless charging vehicles, aerospace, sliding rail equipment and etc [1-3]. In many applications, bidirectional data transmission is necessary in addition to the power transfer. For example, in the condition monitoring of a rotating machinery, the rotating sensor node needs to be powered by the stationary power source. Moreover, the high speed sensing signal and the low speed command signal have to be transmitted between the sensor node and the power controller.

Based on the existing ICPT systems, a low power communication system for embedded devices is designed. The system realizes the efficient and stable wireless power supply. On this basis, a signal transmission system is added, realizing the full duplex real time communication through the mutual coupling coil [4,5]. A frequency selection network is designed to solve the crosstalk problem between power and signal transmission loops.

The design strategy of the inductively coupled power and signal communication system is presented and the optimal element parameters are obtained. Then the feasibility of the scheme is verified using the circuit simulation software. Finally, a hardware experiment platform is built to comprehensively verify the characteristics of power and signal transmission.

II. Full Duplex Power And Communication System

Fig. 1 shows the circuit topology of the system. In power transmission system, after power tube Q_1 to Q_4 of full bridge inverter circuit, the direct current (DC) E1 or E2 is converted into alternating current (AC). By the compensation topology composed of L_1, C_1 and C_2 and coupled coil T0, the AC is transferred to secondary coil. Through compensation capacitor C_3, the load gets power. In the forward signal transmission system, the ICPT system is powered by two DC power sources E1 and E2 with different amplitudes. The switch S is used to control the switching of power supply, so as to realize the modulation of power amplitude. The forward transmitted signal is extracted by a series coupling inductor Ta in the secondary circuit. In the reverse signal transmission system, signal injection and extraction are realized by coupling inductors T1 and T2 in parallel with T0.
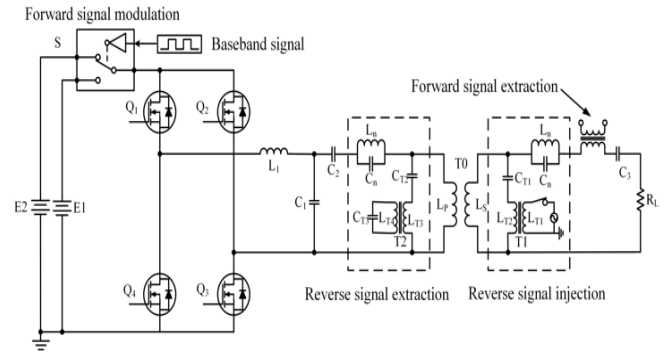
Fig. 1 Circuit topology of the full duplex power and communication system.

In power transmission circuit, L_1, C_1 and C_2 compose a LCC high order compensation topology in primary side. Compared to the low order compensation topology composed of a single capacitor, the high order compensation topology system has a lower sensitivity. The T shape network could achieve an impedance transform and an adjustable output gain. This can realize a constant voltage or current output, which is of great significance in the inductive power and signal transmission system. To realize the full duplex communication, the forward signal transmission system adopts the amplitude modulation (AM) based on electric energy and the reverse transmission employs the high frequency injection technology. In the forward signal transmission system, the carrier signal has a relative low frequency less than 100 kHz. In the reverse signal transmission system, the carrier signal has a frequency higher than 1 MHz. Due to the bandpass and bandstop characteristics of the parallel and serial LC circuits, they are respectively included in the power and signal transmission circuits to realize the isolation between the low frequency and high frequency carrier signals. The crosstalk between the power and signal transmission can be greatly suppressed by using this approach.

III. CIRCUIT CHARACTERISTIC ANALYSIS

A. Power Transmission Circuit Analysis

Through the existing inverter circuit and compensation topology [6-9], a square wave signal is obtained. The electric energy transferred to secondary coil becomes sine form due to the frequency selection property of the resonance circuit. For the convenience of circuit analysis, front DC power and the inverter circuit part can be equivalent to a AC voltage source. U_i is the output peak voltage of AC voltage source, and the equivalent circuit of power transmission system is shown in Fig. 2.
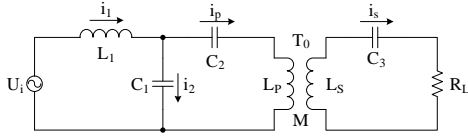
Fig. 2 Power transfer equivalent circuit.

In the power transmission system, ω_0 is the angular frequency of power transmission, M is the mutual inductance between coils, R_L is the equivalent load, and L_S and R_S are the self-inductance and equivalent internal resistance of the secondary coil, respectively.

High order compensation topology LCC can be equivalent to a symmetric T circuit model. According to the characteristics of the symmetric T circuit, it can be obtained as:

$$\begin{cases} jX = j\omega_0 L_1 \\ -jX = \dfrac{1}{j\omega_0 C_1} \\ jX = j(\omega_0 L_P - \dfrac{1}{\omega_0 C_2}) \end{cases} \quad (1)$$

According to the mutual inductance model, the voltage and current equations of the primary and secondary circuit can be obtained as:

$$\begin{cases} \dot{V}_i = \dot{I}_1 Z_P - j\omega_0 M \dot{I}_s \\ -j\omega_0 M \dot{I}_P + \dot{I}_s Z_s = 0 \end{cases} \quad (2)$$

where Z_P is the impedance of the primary circuit, and Z_S is the impedance of the secondary circuit. According to equations (1) and (2), the voltage of the load can be obtained as follows:

$$U_L = \frac{R_L M U_i}{(R_L + R_S) L_1} \quad (3)$$

The output power is:

$$P_o = \frac{M^2 U_i^2 R_L}{(R_L + R_S)^2 L_1^2} \quad (4)$$

The efficiency of power transmission is:

$$\eta = \frac{\omega_0^2 M^2 R_L}{(R_L + R_S)^2 R_P + (R_L + R_S)\omega_0^2 M^2} \quad (5)$$

By ignoring the coil resistance, the LCC-S power transmission system can realize a constant voltage output without changing with the load. Similarly, the LCC high order compensation topology, when equivalent to T-symmetric network, has the characteristics of impedance transformation. The output gain can be flexibly adjusted according to the characteristics of the T circuit, so as to achieve the desired output voltage without introducing additional reactive power.

B. Signal Transmission Circuit Analysis

According to the characteristics of parallel and serial LC circuits, the electrical energy has a large impedance in the signal transmission channel. Hence, the signal transmission circuit can be analyzed independently from the power transmission system when analyzing the reverse signal transmission system. Its equivalent circuit is shown in Fig. 3.
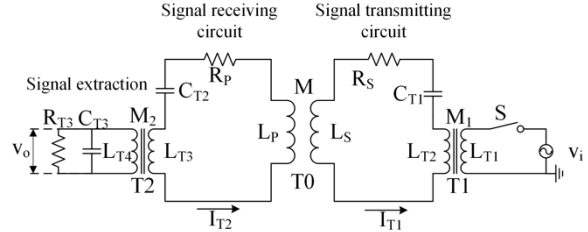
Fig. 3 Reverse signal transmission circuit.

In the parallel resonant circuit composed of L_{T4}, C_{T3} and R_{T3}, the resonant frequency of the circuit is the signal carrier frequency. The equivalent impedance of inductor L_{T4} and capacitor C_{T3} in parallel is very large, which is much larger than the load resistance R_{T3}. The total impedance is determined by the load resistance, so the equivalent impedance of the parallel network is R_{T3}.

The voltage equation is also obtained for the primary and secondary side circuit of the signal transmission loop:

$$\begin{cases} \dot{v}_i = \dot{I}_{T1} Z_{T1} - j\omega_1 M \dot{I}_{T2} \\ -j\omega_1 M \dot{I}_{T1} + \dot{I}_{T2} Z_{T2} = 0 \end{cases} \quad (6)$$

where Z_{T1} is the impedance of the signal sending circuit, and Z_{T2} is the impedance of the signal receiving circuit.

Signal transmission is different from electric energy transmission. In order to ensure the transmission efficiency of electric energy, the compensation network should be added to nullify the consumption of electric energy by the coil self-inductance. However, for the signal transmission, if the electric energy transmission coil resonates with the signal injection and extraction circuit, the impedance of the whole circuit is small and the current passing through is relatively large. Meanwhile, the output voltage of the signal needs a certain time to reach the steady state due to resonance, which seriously affects the demodulation of the signal in high speed signal transmission. Therefore, the self-inductance of the coupling coil is not involved in the calculation of resonance in the design of the signal transmission loop. As a whole, the signal transmission loop presents inductive characteristic.

The equivalent inductance of the secondary signal transmission loop is defined as $L_S{}'$, and the equivalent electric induction meets the following requirement:

$$j\omega_1 L_S{}' = j\omega_1 (L_S + L_{T2}) + \frac{1}{j\omega_1 C_{T1}} \quad (7)$$

Then the signal receiving circuit impedance is:

$$Z_{T1} = R_S + j\omega_1 L_S{}' + \frac{\omega_1^2 M^2}{Z_{T2}} \quad (8)$$

Defining the equivalent inductance of the signal receiving circuit to be $L_P{}'$, the equivalent electric induction meets the following requirement:

$$j\omega_1 L_P{}' = j\omega_1 L_P + \frac{1}{j\omega_1 C_{T2}} \quad (9)$$

The impedance of the signal transmission circuit is:

$$Z_{T2} = j\omega_l L_P' + R_P + R_{T3} \tag{10}$$

According to (6) ~ (10), the extracted signal voltage can be:

$$\dot{v}_o = \frac{j\omega_1 M R_{T3}\left(R_P + R_{T3} + j\omega_l L_P'\right)\dot{v}_i}{\left(R_S + j\omega_l L_S'\right)\left(R_P + R_{T3} + j\omega_l L_P'\right) + \omega_1^2 M^2} \tag{11}$$

The coupling inductors T1 and T2 have the same structure and the coupling coefficients are close to 1. When the signal is transmitted, there is no power loss within the inductors. The transmitting power of the signal can be expressed as

$$P_{i1} = \frac{\left(R_P + R_{T3} + j\omega_l L_P'\right)v_1^2}{\left(R_S + j\omega_l L_S'\right)\left(R_P + R_{T3} + j\omega_l L_P'\right) + \omega_1^2 M^2} \tag{12}$$

Higher frequency of the transmitted signal corresponds to the greater impedance in coil self-induction. As can be seen from the impedance expression of Z_{T2} in the signal transmitting loop, the high frequency of the carrier signal leads to a large impedance of the signal transmitting circuit thus small signal transmitting power. During the signal transmission, partial electric elements of the primary and secondary circuits do not participate in the resonance, resulting into an inductive total impedance and moderate signal transmission power.

IV. SIGNAL MODULATION AND DEMODULATION CIRCUIT DESIGN

A. Forward Signal Modulation Circuit

The key to realize amplitude modulation is how to control the switching of two different amplitude voltages through baseband signal. As the conduction resistance of the switching device needs to be small and the allowable current needs to be relatively large for energy transmission, the traditional integrated signal switch cannot meet the requirements. A switching circuit is designed with the schematic diagram shown in Fig. 4. The power switching circuit of this system is realized by the power MOS tube.
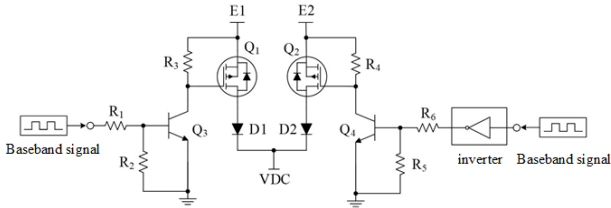
Fig. 4 Amplitude modulation circuit.

In Fig 4, the power MOS tube on two sides control the on and off of E1 and E2 powers, respectively. The working principle of the circuit is as follows. Transistor Q_3 is on when input signal is high. Then P-MOS tube grid voltage is low and the P-MOS tube Q_1 is on. The electric level changes from high to low when the input signal of right side circuit passes the inverter. Transistor Q_4 is off. Q_2 grid voltage is high, so Q_2 is off, too. At this time the system is powered by E1. Similarly, when the signal input is low, Q_1 is off and Q_2 is on, and the system is powered by E2. Diode D_1 and D_2 can isolate the circuit on both sides to prevent current backflow which affects the normal operation of components. The power MOS tube has a small conduction resistance and a high withstand voltage.

Depending on the type, the allowable current of it can reach dozens of ampere. This meets the design requirements.

B. Reverse Signal Modulation Circuit

According to different modulation modes, three digital modulation strategies are most widely used including the Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK) and Phase Shift Keying (PSK). They transmit signals by changing the amplitude, frequency and phase of the carrier signal, respectively.

The realization of FSK modulation requires two high frequency signal sources. The existence of two signal sources increases the complexity of system design, and it is also difficult to realize the continuity of phase during signal switching. PSK modulation need to change the phase with signal phase shifter. It has the strongest anti-interference ability in three kinds of modulation mode, but the coherent demodulation method is needed for the signal demodulation, which demands strict synchronization between output signal and local carrier. Signal transmission exists a time delay in ICPT system, therefore it is difficult to achieve the synchronization between the output signal and the local carrier. In three demodulation methods, it is the most difficult one to achieve. Compared with the previous two methods, the modulation method of ASK is the easiest to realize. Not only the synchronization is not needed and the accuracy requirement of the signal frequency is low, but also the demodulation method is flexible [10]. Therefore, the modulation method of ASK is adopted in the reverse signal transmission system.
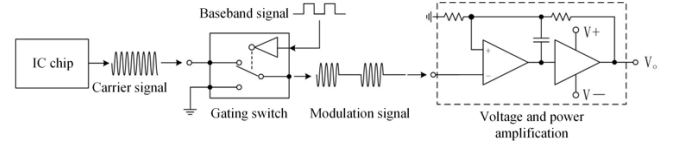
Fig. 5 ASK modulation circuit.

Fig. 5 shows the scheme of the reverse signal modulation circuit based on ASK. Firstly, the high frequency sinusoidal carrier signal is generated through the integrated IC chip. The sinusoidal and ground signal are used as inputs of the dual-channel analog switch. The switch of two signals is controlled by baseband signal. Basing on that the modulation of ASK signal is realized. ASK signal is finally amplified by voltage and power amplifiers to improve the amplitude and load carrying capacity.

C. Demodulation Circuit

The energy amplitude information of the forward signal demodulation can be extracted through the envelope detection. The reverse signal transmission is realized based on the ASK modulation and the demodulation also relies on the envelope detection method.
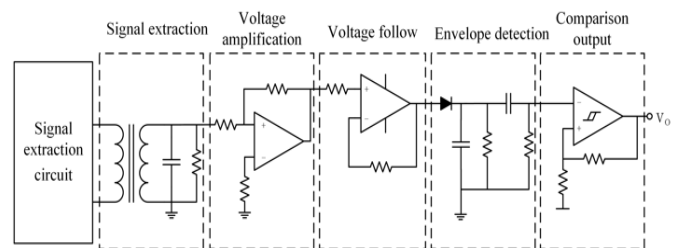
Fig. 6 Signal demodulation circuit.

Fig. 6 is the flow chart of the signal demodulation circuit. High frequency carrier signal is extracted by the coupling inductance. The coupling inductance also plays a role of electrical isolation between the signal extraction circuit and the demodulation circuit. The voltage amplitude extracted by the coupling inductance is small, while the envelope detection circuit is only suitable for demodulation of large signals due to the presence of diode voltage drop. The envelope detection effect is not ideal for signals with amplitude less than 0.5V, so voltage amplification should be carried out after the carrier signal is extracted. Then the voltage follower is used to improve the load carrying capacity, and the diode envelope detection circuit is used to demodulate. Finally the hysteretic comparison circuit is used to achieve the output of the signal.

V. BUILD HARDWARE PLATFORM AND EXPERIMENT ANALYSIS

The experimental platform of the power and signal composite transmission system is shown in Fig. 7. According to calculation and simulation results of the simulation software Multisim, the parameters are shown in TABLE I. The system consists of an inductive power transmission system, a forward signal transmission system and a reverse signal transmission system. The power transmission system mainly consists of a DC voltage regulator circuit, an inverter circuit, a LCC-S compensation circuit, a planar coil coupling mechanism, a rectifier and an energy storage circuit. The wireless transmission distance of the power transmission system is 5 cm. The forward signal transmission system mainly consists of power switching circuit and forward signal demodulation circuit. The reverse signal transmission system consists of ASK modulation circuit and incoherent demodulation circuit.
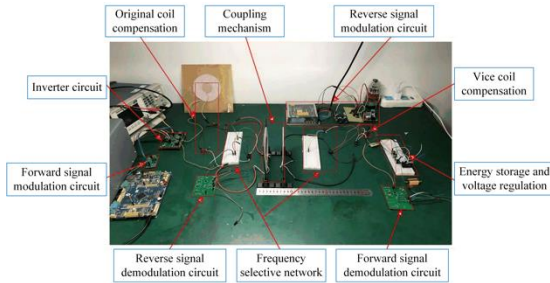
Fig. 7 Experimental platform of power and signal transmission system.

TABLE I. REVERSE SIGNAL TRANSMISSION SIMULATION PARAMETERS

Parameters	Value
L_n	10uH
C_n	6.33nF
C_{T1}	1.408nF
L_{T1}, L_{T2}	4.5uH
L_{T3}, L_{T4}	4.5uH
v_i	5V
f_1	2MHz
M_1	4.2uH
M_2	4.2uH
C_{T3}	1.408nF

L_n, C_n, C_{T1} are inductor and capacitors of the frequency selective network. L_{T1}, L_{T2} are signal injection coil self-inductance. L_{T3}, L_{T4} are signal extraction coil self-inductance. v_i, f_1 are voltage and frequency of high frequency carrier signal. M_1 is signal injection coil mutual inductance. M_2 is

signal extraction coil mutual inductance and C_{T3} is signal extraction compensation capacitor.

In power transmission experiment, DC input voltage is 7V, inverter frequency is 50kHz, and equivalent load impedance is 20Ω. Fig. 8 shows the experimental waveform of power transmission. It can be seen from the experiment that the inverter bridge outputs the square wave signal after DC passes through the inverter circuit. The peak interference in the square wave is small, indicating that the impact received by the inverter circuit is very small when switch works. It can be seen from the voltage waveform of primary side coil that the high harmonic component of the square wave is well suppressed in the primary side LCC resonant network, and only the sinusoidal power waveform exists. The peak voltage of the primary side coil reaches 38 V, and the peak voltage transferred to secondary coil reaches 32 V. Compared with the input 7 V DC voltage, the resonant network gives a certain gain in voltage amplitude, which is of great significance in the power transmission system with a low coupling coefficient. The AC power waveform transferred to the secondary is stable at 3.3 V after the full-bridge rectifier and the voltage regulator circuit, which can directly supply power to embedded devices. It can be seen from the voltage waveform after voltage regulator, the ripple of the output voltage is very small, and the voltage amplitude is stable, realizing the function of rectification and voltage regulation.
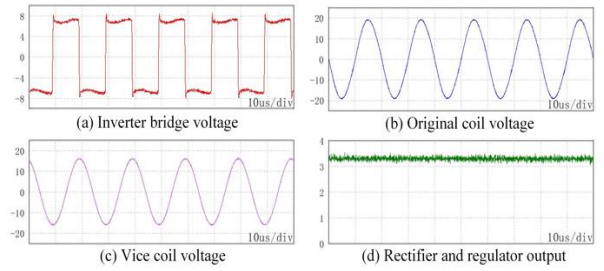
Fig. 8 Experimental waveform of power transmission.

In the forward signal transmission system, input DC power amplitude are 7 V and 10 V. The transmitting baseband signal is generated by a single-chip microcomputer, and the power amplitude modulation is realized by controlling the power switching circuit. In the experiment, the baseband signal is simulated by the square wave signal produced by the single chip microcomputer. High level output signal is '1', and low level is '0'. Fig. 9 shows the experimental waveform of forward signal modulation and demodulation. In Fig. 9, the modulated waveform is the energy waveform extracted from the coupling inductance in series on the secondary loop. The amplitude extracted by the signal is smaller than the voltage of the secondary coupling coil because the self-inductance of the coupling inductance is far less than that of the power transmission coil. From the modulation results, it can be seen that the amplitude of the electric energy becomes stable in a very short time, which realizes a fast amplitude modulation. In the switching circuit that the electric energy amplitude changes, there is no energy storage element such as inductor or capacitor, and no charging and discharging process of energy storage elements when the electric energy amplitude is switched, so the response time is very short. The high and low amplitude of power amplitude modulation is obvious, and it can be precisely realized in the subsequent envelope detection and hysteresis comparison circuit. Finally, the primary baseband signal is restored completely.
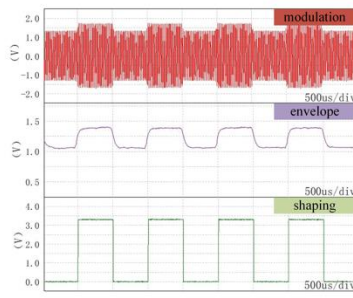
Fig. 9 Forward signal transmission waveform.

Fig. 10 shows the waveform in reverse signal transmission experiment. The carrier signal is loaded into the power transmission circuit after ASK modulation. It can be seen from the waveform of two ends of secondary coil. As the signal amplitude is far less than the amplitude of electric energy, the signal noise ratio (SNR) of signal carrier in electric transmission circuit is low and the reverse signal transmission will not affect the power transmission. In the signal receiving circuit, due to the presence of frequency selecting network, the low-frequency power wave has a great impedance in the signal transmission circuit, resulting into almost no interference of the electric energy in the signal extraction waveform. The modulation of the reverse signal is consistent with the simulation results. High frequency signal has a great impedance in the power transmission circuit, so the extracted signal amplitude is only about 0.8V, which need be amplified by signal amplification circuit. The magnification is 4 times in signal amplifier circuit of demodulation. The amplified signal provides the input for the envelope detection and hysteresis comparator modules. Signal demodulation can be accurately realized due to the high SNR.
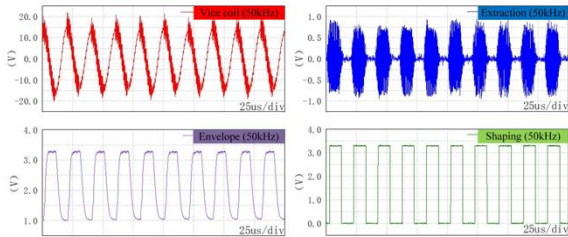
Fig. 10 Reverse signal transmission waveform.

The parallel LC circuit have a great impedance to the high frequency carrier signal. Fig. 11 shows the voltage of two ends of secondary coil and power waveform after parallel LC circuit. It can be seen that the frequency selective network can clearly select between different frequency waveforms. The subsequent channel of power transmission system is not interfered by the high frequency carrier signal. The existence of the frequency selective network effectively ensure the normal operation of the signal and power transmission.
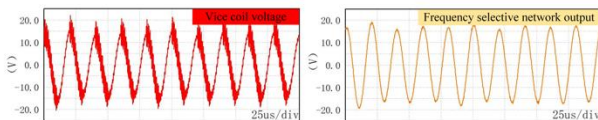
Fig. 11 Filtering characteristics of frequency selective network.

Fig. 12 shows the experimental waveform of full duplex communication of the system. The forward signal transmission rate is 2 kbps and the reverse signal transmission rate is 150 kbps. When the high frequency carrier signal is loaded into the secondary coil, the forward power modulation waveform is not very clear due to the influence of high frequency carrier signal. However, the forward modulation waveform is clear due to the forward signal extraction circuit after the frequency selective network. The normal demodulation can be realized. Similarly, the reverse signal transmission channel has a impedance on the low frequency electric energy, and the change of the electric energy amplitude will not affect the modulation and demodulation of the reverse signal. To sum up, due to the different modulation principle and the existence of frequency selective network, there is no interference between forward signal transmission and reverse signal transmission. The full duplex communication of the system can be realized.
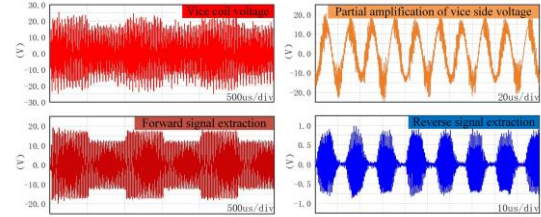
Fig. 12 Full duplex signal transmission waveform.

VI. CONCLUSION

In this paper, a LCC-S type power transmission system is designed, and the theoretical model and power transmission characteristics of the system are analyzed in detail. In addition, aiming at overcoming the deficiency of the existing power transmission system, an energy storage and power supply circuit is designed at the power receiving end to realize the self-power supply function of the system. The power of energy transmission in the system can reach up to 5 W, and the transmission efficiency can reach 75% under optimal load. Based on the power transmission system, a full duplex communication system is designed with the power wave as the transmission medium. In the forward signal transmission system, the maximum signal transmission rate is 10 kbps. And in the reverse communication system, the maximum signal transmission rate is 200 kbps. In the communication system, the anti-interference performance between signals and the influence on power transmission are analyzed.

REFERENCES

[1] Y. Sun, P. X. Yan, Z. H. Wang, and Y. Y. Luan, "The parallel transmission of power and data with the shared channel for an inductive power transfer system," IEEE Trans. Power Electron., vol. 31, no. 8, pp. 5495–5502, Aug. 2016.

[2] C. C. Huang, C. L. Lin, and Y. K.Wu, "Simultaneous wireless power/data transfer for electric vehicle charging," IEEE Trans. Ind. Electron., vol. 64, no. 1, pp. 682–690, Jan. 2017.

[3] Simard G, Sawan M, Massicotte D, et al, "High-Speed OQPSK and Efficient Power Transfer Through Inductive Link for Biomedical Implants," IEEE Transactions on Biomedical Circuits and Systems, vol. 4, no. 3, pp. 192-200, Jun. 2010.

[4] J. Wu, C. Zhao, Z. Lin, J. Du, Y. Hu and X. He, "Wireless Power and Data Transfer via a Common Inductive Link Using Frequency Division Multiplexing," in IEEE Transactions on Industrial Electronics, vol. 62, no. 12, pp. 7810-7820, Dec. 2015.

[5] C. Li, Z. Wang, Y. Sun and X. Dai, "Full Duplex Communication Based on Partial Power Coil in Inductive Coupling Power Transfer System," 2019 IEEE PELS Workshop on Emerging Technologies: Wireless Power Transfer (WoW), London, United Kingdom, pp. 351-356, 2019.

[6] Surajit Das Barman, Ahmed Wasif Reza, Narendra Kumar, Md. Ershadul Karim, Abu Bakar Munir, "Wireless powering by magnetic resonant coupling: Recent trends in wireless power transfer system an

its applications," Renewable and Sustainable Energy Reviews, Elsevier, vol. 51(C), pp. 1525-1552, Jul. 2015.

[7] F. Lu, H. Zhang, H. Health, and C. C. Mi, "An inductive and capacitive combined wireless power transfer system with LC-compensated topology," IEEE Trans. Power Electron., vol. 31, no. 12, pp. 8471–8482, Dec. 2016.

[8] Li X, Tang C, Dai X, et al, "An Inductive and Capacitive Combined Parallel Transmission of Power and Data for Wireless Power Transfer Systems," IEEE Transactions on Power Electronics, vol. 33, no. 6, pp. 4980-4991, Jun. 2018.

[9] C. Xia, W. Wang, G. Chen, X. Wu, S. Zhou and Y. Sun, "Robust Control for the Relay ICPT System Under External Disturbance and Parametric Uncertainty," IEEE Transactions on Control Systems Technology, vol. 25, no. 6, pp. 2168-2175, Nov. 2017.

[10] Qingxin Yang et al, "Optimal design of energy transmission system for implantable device base on WiTricity," Digests of the 2010 14th Biennial IEEE Conference on Electromagnetic Field Computation, Chicago, IL, pp. 1-1, 2010.

Ultra-micro Vibration Measurement Method Using CW Doppler Radar

Songxu Li
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
songxu.li@sjtu.edu.cn

Yuyong Xiong
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
yy,xiong@sjtu.edu.cn

Zesheng Ren
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
Silence-SCIENCE@sjtu.edu.cn

Changzhan Gu
MoE Key Laboratory of Design and Electromagnetic Compatibility
of High Speed Electronic System
Shanghai Jiao Tong University
Shanghai, China
changzhan@sjtu.edu.cn

Zhike Peng*
State Key Laboratory of Mechanical System and Vibration
Shanghai Jiao Tong University
Shanghai, China
z.peng@sjtu.edu.cn

Abstract—As a non-contact vibration measurement technology, microwave radar is widely used in vital sign monitoring, structure health monitoring and precision instrument measurement. Although significant progress has been made, the continuous wave (CW) Doppler radar has inherent defects in ultra-micro vibration measurement due to the DC offset problem, which is challenging to be solved by the common circle fitting method with orthogonal baseband signals. In this paper, based on the Doppler phase demodulation algorithms, a pre-calculation method, creating a long arc of circle fitting by sliding the radar front-end in advanced, is proposed to accurately estimate the DC offset. Theoretical analysis and experimental results were provided for demonstrating our proposed method, which show that the ultra-micro vibration measurement of 5μm can be realized.

Keywords—*CW Doppler radar, pre-calculation, DC offset, ultra-micro vibration, vibration measurement*

I. INTRODUCTION

The vibration measurement based on microwave sensing has obvious advantages compared with the traditional contact sensors, which has been extensively studied in vital sign monitoring [1], large-scale structure state monitoring [2] and precision mechanical measurement. And the displacement can be obtained by phase demodulation from the baseband signal which is obtained by mixing the received signal and the transmitted signal of continuous wave (CW) Doppler radar. In order to avoid the null point detection problem happened when the phase related to the distance between the radar and the target is a multiple of π, the quadrature receiver architecture of CW Doppler radar has been widely investigated.

Various methods have been proposed for nonlinear phase demodulation of quadrature CW radar, such as the Arctangent demodulation (AT) [3], complex signal demodulation (CSD) [4] and the parameterized demodulation (PD) [5]. However due to the defects of the radar system and static targets in the measurement environment, the I/Q imbalance correction and DC offset compensation must be implemented before calculating the displacement. Singh et al. presented a data-based quadrature imbalance compensation method, which is demonstrated as an efficient and accurate imbalance correction method [6]. Park et al. presented a center estimation method for DC offset compensation. However, it requires a long enough arc to fit the exact center of the circle, and it is not suitable for the measurement of ultra-micro motions.

For the purpose of solving the problem of ultra-micro amplitude vibration measurement, RF phase shifter can make the arc long enough to fit into a circle, but the high-frequency phase shifter is too expensive to apply on the radar [7]. Kim et al. estimated the center of small vibration displacement through the Quadrature Frequency-Group (QFG) radar [8], while the center of the circle changes with the change of the center frequency of the microwave signal. Huang et al. proposed a semi-definite programming algorithm which requires complex calculation [9].

In the view of the challenge of ultra-micro amplitude vibration measurement, a method of moving the front end of the radar to obtain long enough arc is proposed in this paper. In section II, a brief introduction of the ultra-micro amplitude vibration measurement theory is given. Section III presents the experiments and the analysis results. A summary of the contribution of this paper is provided in Section IV.

II. THEORY
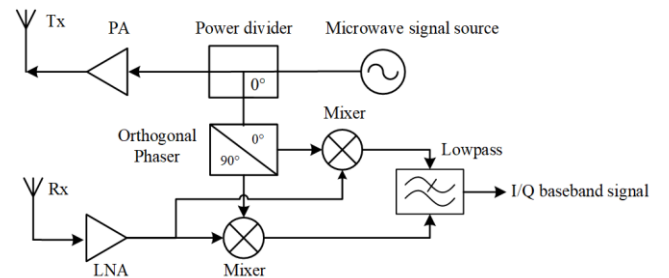
Fig. 1. The block diagram of the quadrature CW Doppler radar system.

Fig. 1 shows the block diagram of the quadrature CW Doppler radar system. The microwave signal source transmitted by the power divider is sent out through the power amplifier (PA) and the transmitting antenna, and the local oscillator (LO) signals of I and Q channels are generated by the orthogonal phaser. The receiving signal is connected with the low noise amplifier (LNA), and then the amplified signal is mixed with the orthogonal signals through the mixer. Finally, the I/Q baseband signal is obtained from the mixing signal by lowpass filtering. And the I/Q baseband signal can be expressed as

$$I(t) = DC_I + A_I \cos(\frac{4\pi x(t)}{\lambda} + \phi_I) \qquad (1)$$

$$Q(t) = DC_Q + A_Q \sin(\frac{4\pi x(t)}{\lambda} + \phi_Q) \qquad (2)$$

where DC_I and DC_Q represent DC offset, the amplitude imbalance is $A_e = (A_Q/A_I)$,the phase imbalance is $\phi_e = \phi_Q - \phi_I$, λ is the wavelength of the microwave signal, and $x(t)$ is the vibration displacement of the target. Firstly, the amplitude imbalance and the phase imbalance should be estimated and compensated before the phase demodulation, which caused by the imperfections of the radar system. Thus the Gram-Schmidt orthonormalization is applied to estimate the amplitude imbalance and the phase imbalance.
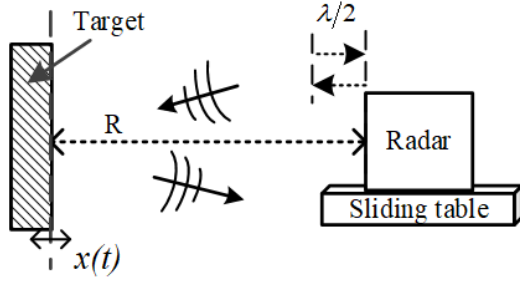
Fig. 2. Demonstration diagram of pre-calibration method.

$$[I_c(t) \quad Q_c(t)] = [I(t) \quad Q(t)] \begin{bmatrix} A_e & 0 \\ -\sin(\phi_e) & \cos(\phi_e) \end{bmatrix} (3)$$

And then the compensated baseband signals can be represented by

$$I_c(t) = DC_I + \cos(\frac{4\pi x(t)}{\lambda} + \theta) \qquad (4)$$

$$Q_c(t) = DC_Q + \sin(\frac{4\pi x(t)}{\lambda} + \theta) \qquad (5)$$
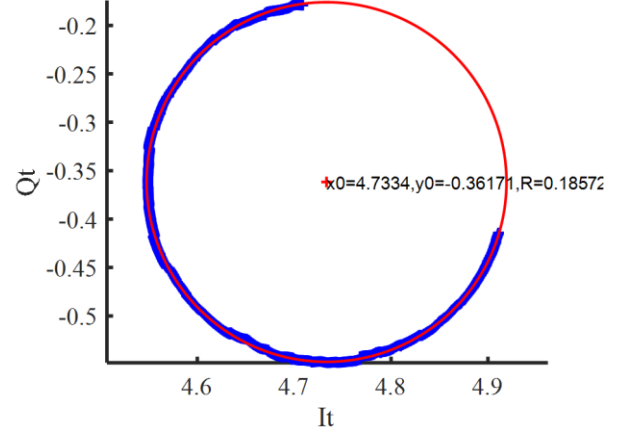
Fig. 3. DC offset estimation by fitting the center of the circle with the least square method.

It can be seen that the DC offset is unavoidable because of the leakage of microwave signals, the self-mixing of the
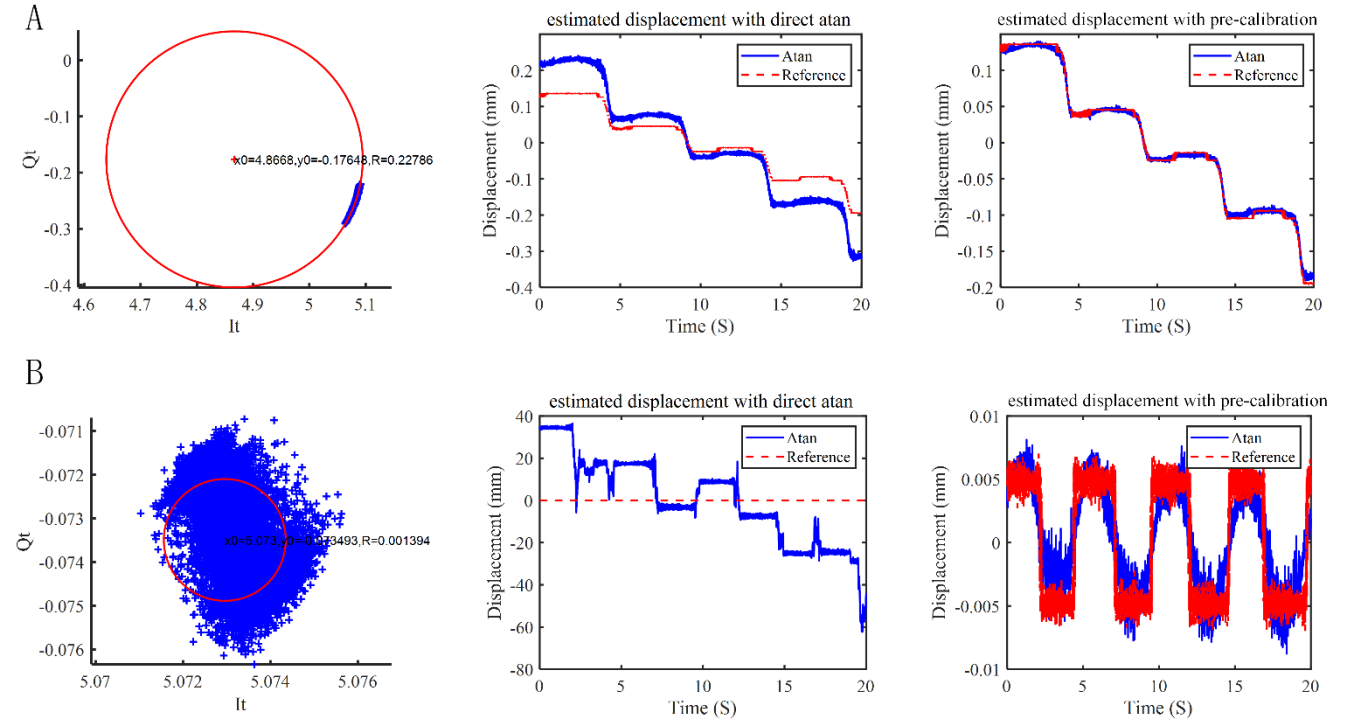
Fig. 4. (A) 100μm amplitude measurement: the left picture shows the center fitting for DC offset estimation, the middle picture shows the result of traditional method with DC offset compensation, and the right picture shows the result obtained by pre-calibration method. (B) 5μm amplitude measurement: the left picture shows the center fitting for DC offset estimation, the middle picture shows the result of traditional method with DC offset compensation, and the right picture shows the result obtained by pre-calibration method.

receiving signals and the static targets in the measurement environment. Here we utilize the least square method for circle fitting to estimate the DC offset DC_I and DC_Q. After eliminating the DC offset, inverse tangent demodulation method is used to obtain the phase of the baseband. And we can easily get the vibration displacement which is positively related to the phase $\varphi(t)$.

$$\varphi(t) = \arctan \frac{Q_c(t) - DC_Q}{I_c(t) - DC_I} \quad (6)$$

$$x(t) = \frac{\lambda}{4\pi}\left[\varphi(t) - mean(\varphi(t))\right] \quad (7)$$

However, the phase change of the baseband signal is insufficient to fit a complete arc when the displacement variation is very small which reaches to dozens of microns, resulting in a large deviation in the center estimation. Here an ultra-micro vibration measurement method is proposed by moving the radar front-end to increase the arc. As shown in Fig. 2, the radar is fixed on a sliding table, and the target is in front of the radar at the distance of R. In order to get a sufficient arc, we make the front end of the radar move $\lambda/2$ back and forth by controlling the sliding table. And the phase of the baseband signal changes from 0 to 2π, which will provide a better circle fitting result for DC offset. After that, we fix the radar at a distance of R from the target to be measured, and collect the microwave radar baseband signals synchronously. After the amplitude compensation and phase compensation of the baseband signal, we can directly calibrate the DC offset which has been estimated during the sliding process. Finally, we can obtain the accurate ultra-micro amplitude vibration measurement results after phase demodulation.

III. Experimental Results

In order to verify the feasibility and effectiveness of our proposed method, we have built a vibration monitoring system based on CW Doppler radar and carried out a series of experiments. The carrier frequency of the microwave signals is 24.26GHz, and the radar transmission power is 20dBm. The beamwidth of the antenna is 11° in both horizontal and vertical directions.

For the purpose of fitting a long enough arc, we control the radar front-end sliding 6 mm (approximately the length of $\lambda/2$) back and forth. After eliminating the amplitude imbalance and phase imbalance, the arc is gained by least square fitting, as shown in Fig. 3. And we can record the value of DC offset for subsequent analysis. The following is the measurement of ultra-micro motions.

Primarily, we set the amplitude of the target to about 100μm, and measure the movement of the target with laser displacement sensor and microwave radar. As shown in Fig. 4A, the left picture is the result calculated by traditional DC offset compensation algorithm. And it is obviously that the small motion can only fit a short arc, which leads to a large deviation in the center estimation. The final vibration displacement measurement result in the middle represents that it can only reflect the general movement trend, but a large displacement error. The same is compensated by the pre-calibrated DC offset, and the measurement result matches well with the reference signal.

To demonstrate the superiority of the pre-calibration method in ultra-micro amplitude measurement, we set the target motion amplitude to 5μm. It can be seen that the baseband signal can only fit a point cloud in Fig. 4B, but not an effective circle. And the displacement is far beyond from the reference. Apparently, the method proposed in this paper can satisfy the motion monitoring as small as 5μm from the right picture of Fig. 4B.
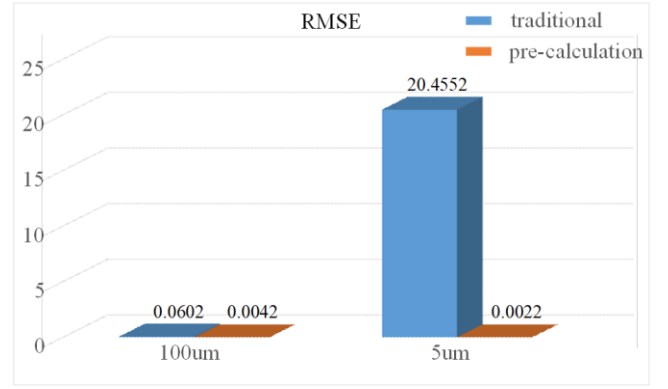
Fig. 5. The RMSE results of the 100μm and 5μm amplitude vibration.

IV. Conclusions

It is difficult to estimate the accurate center of the circle using the least square method when the vibration amplitude is small. Therefore, a pre-calibration method is proposed in this paper, in which the DC offset is obtained by fitting a long arc of the sliding radar front-end. And then the calculated DC offset is used to compensate the baseband signal of ultra-micro motions. The root mean square error (RMSE) of the measurements as shown in Fig. 5, the proposed method can obviously improve the measurement accuracy and will paly an important role in vital sign monitoring and ultra-micro vibration measurement of other precision instruments.

References

[1] Y. Xiong, Peng Z, Gu C, Li S, Wang D, Zhang W, "Differential Enhancement Method for Robust and Accurate Heart Rate Monitoring via Microwave Vital Sign Sensing," IEEE Transactions on Instrumentation and Measurement. 2020.

[2] Zhao W, Zhang G, Zhang J, "Cable force estimation of a long-span cable-stayed bridge with microwave interferometric radar," Computer-Aided Civil and Infrastructure Engineering. 2020.

[3] Park, B.K., Boric-Lubecke, O., Lubecke, V.M., "Arctangent Demodulation With DC Offset Compensation in Quadrature Doppler Radar Receiver Systems," IEEE Transactions on Microwave Theory and Techniques, 2007, 55(5):1073-1079.

[4] Li, C., Lin, J., "Complex signal demodulation and random body movement cancellation techniques for non-contact vital sign detection," Microwave Symposium Digest, 2008 IEEE MTT-S International, 2008:567-570.

[5] Y. Xiong, S. Chen, X. Dong, Z. Peng, and W. Zhang, "Accurate Measurement in Doppler Radar Vital Sign Detection Based on Parameterized Demodulation," IEEE Transactions on Microwave Theory and Techniques, 2017, 65(11): 4483-492.

[6] Singh A, Gao X, Yavari E, Zakrzewski M, Cao XH, Lubecke VM, et al., "Data-Based Quadrature Imbalance Compensation for a CW Doppler Radar System," IEEE Transactions on Microwave Theory and Techniques. 2013;61(4):1718-24.

[7] Xiaomeng, G., Xu, J., Rahman, A., Lubecke, V. and Boric-Lubecke, O., "Arc shifting method for small displacement measurement with quadrature CW doppler radar," Microwave Symposium (IMS), 2017 IEEE MTT-S International, pp. 1003–1006.

[8] Kim DK, Kim Y., "Quadrature Frequency-Group Radar and its center estimation algorithms for small Vibrational Displacement," Scientific Reports, 2019;9(1):6763.

[9] Huang, M. et al. A self-calibrating radar sensor system for measuring vital signs. IEEE Transactions on Biomedical Circuits and Systems, 2016, pp.352–363.

A Selective Anchor Node Method Based on Tetrahedral Volume for Three- dimensional DFL Node Self-localization

BingHua Huang
School of Mechanical Engineering
NanJing University of Science and Technology
NanJing, China
huangbinghua2018@njust.edu.cn

XuLiang Qin
School of Mechanical Engineering
NanJing University of Science and Technology
NanJing, China
njustqxl@163.com

JiaXing Yang
School of Mechanical Engineering
NanJing University of Science and Technology
NanJing, China
j.x.yang@njust.edu.cn

Manyi Wang
School of Mechanical Engineering
NanJing University of Science and Technology
NanJing, China
manyi.wang@njust.edu.cn

Abstract—**Device free localization (DFL) has become one of the promising techniques for object detection, positioning and tracking. The DFL system does not require any electronic device to be attached on the target. It relies the variance of the received signal strength (RSS) of the communication link between every two sensor nodes to locate the target in the monitoring area. Thus, the localization of each sensor node has tight relationship with the accuracy of the DFL. However, the state of the art always assume that the localization of all sensor node has been known beforehand and rely on measuring the coordinate of the sensor nodes manually one by one, which is a very time-consuming and manpower intensive job. In this paper, we propose a self-localization method for sensor nodes in the scenario of three-dimensional DFL. Specifically, we utilize the tetrahedral volume method to select the optimal anchor nodes, and then applying the maximum likelihood estimation algorithm to locate the sensor node. Field experiments shows that the three-dimensional DFL node self-localization technique can locate the sensor node precisely.**

Keywords—self-localization, device free localization, three-dimensional, RSS,wireless sensor networks

I. INTRODUCTION

DFL [1], performed on the WSN platform,is a system that used to detect, track and identify the object without any device attached as well as not actively participate in the localization process. By extracting the various of the RSSI introduced by the object, DFL can not only estimate the location, but also calculate the approximate scale of the object. Compared with other active location methods (e.g. visual positioning, ultra-wideband, infrared imaging technology), DFL can work in the scenarios of darkness, dense fog, obstacle. In addition, DFL system is a cost-sensitive positioning technology relying on simple wireless sensor nodes and has wide range of applications, such as fall detection, roadside surveillance, ambient assisted living and the illegal abroad crossing surveillance in a low-coat way [2-5].

However, the existing research of DFL mostly assume that the sensor node coordinates are known beforehand, which directly leads to a lot of time and manpower to measure the sensor node coordinates before DFL positioning. As a result,

the practical applications of DFL that require emergency deployment and rapid positioning, such as battlefield vehicle perception and anti-terrorism rescue will be limited. To this end, we propose a selective anchor node method based on tetrahedral volume, which can automatically achieve the coordinate acquisition of DFL wireless sensor nodes with high precision, so that the application of DFL for emergency monitoring and disaster relief has been further expanded.

The most popular node positioning algorithms is Global Positioning System (GPS) which is expensive, and it is not suitable for the case where the node volume is too small to carry a GPS equipment. It becomes an important content that how to deploy all the sensor nodes low-cost, quickly and accurately. Therefore, the research on node self-localization possess important research value and practical significance.

The rest of the paper is organized as follows. Section II introduces the related works. Section III proposes the node self-localization algorithm based on the tetrahedral volume. Section IV validates the proposed model and system with extensive experiments. Finally, we make a conclusion in Section V.

II. RELATED WORKS

There are many investigation papers in the node positioning field. Nowadays, node positioning algorithms can be mainly divided into ranging-based and non-ranging-based positioning algorithms. Common measurement methods based on non-ranging positioning contains: distance vector-Hop (DV-Hop) [6], centroid, approximate point-in-triangulation test (APIT), multi-dimensional scaling (MDS), which have few requirements on hardware while the accuracy is bad. On the contrary, methods based on ranging like time of arrival (TOA) [7], time difference of arrival (TDOA), angle of arrival (AOA), and received signal strength indicator (RSSI), have relatively high accuracy and high requirements on hardware. Among them, the RSSI-based method has the characteristics of low cost, low power consumption, and strong practicability compared with the other methods [8][9]

Therefore, according to the characteristics of the DFL system, a positioning method based on RSSI is used in this work.

There are lots of researches about the node self-localization for WSNs. Kumar et al. [10] proposed a self-localization method which is based on RSSI with perturbed anchor positions. Abouzar et al. [11] proposed a RSSI-based distributed Bayesian localization algorithm based on message passing to solve the approximate inference problem. Those methods are all focus on two-dimensional nodes self-localization. To solve three-dimensional problems, Rui et al. [12] used weight matrix to improve the maximum likelihood estimation positioning algorithm. In addition, Kirichek [13] developed a method of rotating coordinate axis in order to avoid solving quadratic equation. However, those approaches do not take into account the topology of the nodes which is a significant issue to impact the positioning effect.

By analyzing the previous work, we propose a novel node self-localization for three-dimensional DFL system by selecting anchor node based on tetrahedral volume, which can locate the nodes automatically, cheaply and precisely.

III. THEORETICAL ANALYSIS OF THIS WORK

In this section, we present the node self-positioning method and give a theoretical analysis. The process of the method is as follows:

First of all, we choose a measurement model to determine the distance between the sending and the receiving node with the least error. Then, we utilize the maximum likelihood estimation to calculate the three-dimensional coordinate. Finally, we propose the selective anchor node method based on tetrahedral volume to choose the most proper anchor nodes. The node positioning flowchart is shown in Figure 1.

A. Distance Measurement Model Based on RSSI

RSSI is an estimated measure of power level that sensor nodes received from others. Specifically, it means that when a certain frequency signal propagates in space, the wireless signal attenuation increases gradually with the distance increasing, leading to the received radio signal power decreases. Using a fixed transmit frequency, a corresponding relationship is existed between RSSI and distance. According to this relationship, the distance between nodes can be determined once the RSSI between nodes is obtained.

According to related works, commonly used ranging models include log-normal distribution empirical models and polynomial models.

The log-normal distribution empirical model is based on the free space propagation model, taking the logarithms on both sides, andthe model can be expressed as:

$$RSSI\left(d\right) = RSSI\left(d_0\right) + 10nlog\left(\frac{d}{d_0}\right) + A + e \quad (1)$$

where d denotes the distance between the sending and the receiving nodes, $RSSI(d)$ denotes the received power, $d0$ and $RSSI(d_0)$ denotes the coefficients of the reference point, n denotes the path loss which is affected by the environmental factors, A denotes the attenuation correction term which is equal to 0 when attenuation effects are ignored, ε denotes the error correction term which follows a normal distribution with 0 as the mean.
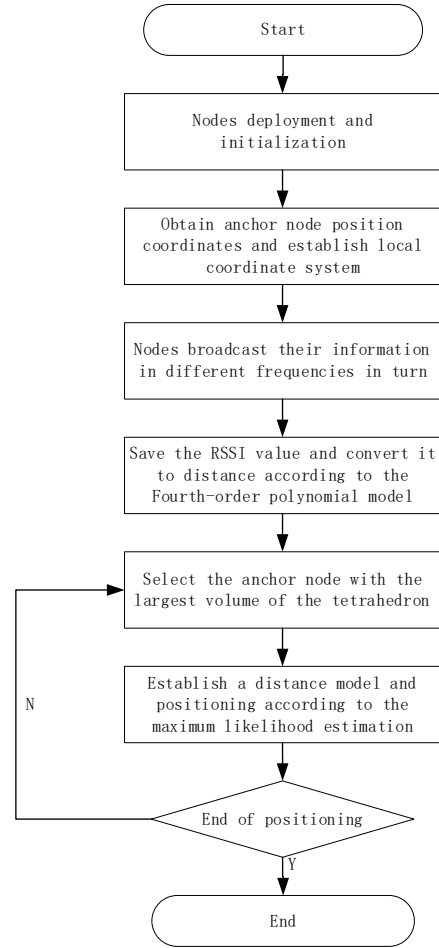
Figure. 1 . Node positioning flowchart

A polynomial model can be directly used to fit the relationship between the distance and the RSSI. There are third-order or higher-order polynomial models we can choose. No matter which model is used, the parameters need to be calibrated before the experiment using the least square method.

Since there are many models used to fit the relationship between RSSI and distance, under-fitting and over-fitting may occur. In order to selected the most suitable model among these as the subsequent measurement of the distance between nodes, the standard need to be set.

The root mean square error (*RMSE*) can well represent the precision of the model. The model with the smallest *RMSE* is the best. The root mean square error expression is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(d_{real} - d_{cal}\right)^2}{n}} \quad (2)$$

where d_{real} denote the real distance between nodes, d_{cal} denotes the calculated distance between nodes according to the model, and n is the number of observations.

B. Maximum Likelihood Stimation

In the WSNs nodes self-localization system, the number of anchor nodes is n, the coordinates are (x_1, y_1, z_1), (x_2, y_2, z_2) ⋯ (x_n, y_n, z_n). Assuming that the coordinate of them is (x_u, y_u, z_u), and the distance between nodes i and j is d_{ij}. We can describe the relationship between the nodes by establishing a series of equations using the distance equations in 3D space:

$$\begin{cases} (x_1 - x_u)^2 + (y_1 - y_u)^2 + (z_1 - z_u)^2 = d_{1u}^2 \\ (x_2 - x_u)^2 + (y_2 - x_u)^2 + (z_2 - z_u)^2 = d_{2u}^2 \\ \vdots \\ (x_n - x_u)^2 + (y_n - x_u)^2 + (z_n - z_u)^2 = d_{nu}^2 \end{cases} \quad (3)$$

The formula (3) is a system of quadratic equations. It is not computable directly using a computer. So we start from the first formula successively subtracting the last equation, then obtain the position of the node as the linear equation (4) express.

$$X = \left(A^T A \right)^{-1} A^T b \quad (4)$$

Where X, A, b can be represented by the matrix:

$$X = \begin{bmatrix} x_u \\ y_u \\ z_u \end{bmatrix} \quad (5)$$

$$b = \begin{bmatrix} x_n^2 - x_1^2 + y_n^2 - y_1^2 + z_n^2 - z_1^2 + d_{1u}^2 - x_{nu}^2 \\ x_n^2 - x_2^2 + y_n^2 - y_2^2 + z_n^2 - z_2^2 + d_{2u}^2 - x_{nu}^2 \\ \vdots \\ x_n^2 - x_{n-1}^2 + y_n^2 - y_{n-1}^2 + z_n^2 - z_{n-1}^2 + d_{(n-1)u}^2 - x_{nu}^2 \end{bmatrix} \quad (6)$$

$$A = \begin{bmatrix} 2(x_1 - x_u) & 2(y_1 - y_u) & 2(z_1 - z_u) \\ 2(x_2 - x_u) & 2(y_2 - y_u) & 2(z_2 - z_u) \\ \vdots \\ 2(x_{n-1} - x_u) & 2(y_{n-1} - y_u) & 2(z_{n-1} - z_u) \end{bmatrix} \quad (7)$$

C. Selecting the Anchor Nodes Baseon the Tetrahedral Volume

In the three-dimensional node self-positioning, at least four anchor nodes are required to colete the positioning of the target node. These four anchor nodes are form a positioning unit in space. The topological quality of the positioning unit greatly affects the positioning accuracy.

When the volume of the positioning unit is very small, the error will be very large, which is called the occurrence of multicollinearity. In [14][15], the tetrahedron composed by anchor nodes has nine types and the topological quality is poor. Among all these nine kinds of poor-quality tetrahedron composed by anchor nodes, the volume is small. Therefore, in order to improve the positioning accuracy, it is necessary to calculate the volume of tetrahedron and find out the anchor node combined with the largest volume.

Assuming that the four anchor nodes are P_a, P_b, P_c and P_d respectively, according to the vector mixed product formula, calculate the volume of the tetrahedron. Choosing P_a as the vertex of the tetrahedron, the three vectors passing through the vertex are: $v_1 = P_b - P_a$, $v_2 = P_c - P_a$, $v_3 = P_d - P_a$. Then the volume of the tetrahedron is:

$$V = \left| (v_1 \times v_2) g v_3 \right| / 6 \quad (8)$$

Calculated with the determinant:

$$V = \frac{1}{6} \times \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \end{vmatrix} \quad (9)$$

Among this anchor nodes, find the tetrahedron with the largest volume and return the corresponding anchor node numbers.

IV. EXPERIMENTS

In order to verify the improvement of the aforementioned algorithm for the positioning accuracy, the following experiments are implemented: The ZigBee/CC2530 programmable hardware wireless modules are adopted as wireless nodes to build the three-dimensional WSN system as shown in Fig. 2. Before the experiments, the parameters in distance and RSSI expressions are calibrated by using different models to fit the distance between nodes and RSSI, and the results obtained are shown in Table I.
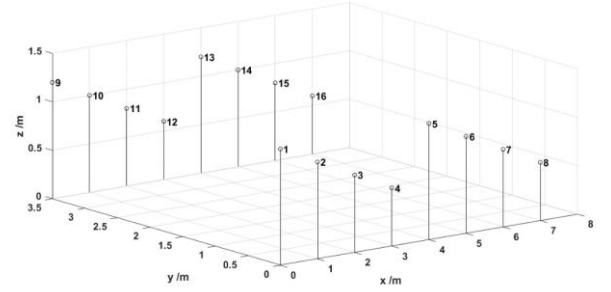
g. 2. Sensor node layout

As shown in Table I, among the four models, the fourth-order polynomial model has the smallest RMSE being 1.6m, which is 49.8% more accurate than the log-normal empirical model.

TABLE I MODELS AND RMSE

Model	Mathematical expression	RMSE
Log-normal empirical model	$RSSI = -34.95 - 8.31 \log(d)$	3.1m
Third-order polynomial model	$RSSI = -0.14d^3 + 2.24d^2 - 12.50d - 23.75$	2.9m
Fourth-order polynomial model	$RSSI = 0.05d^4 - 0.91d^3 + 6.44d^2 - 20.87d - 19.30$	1.6m
Fifth-order polynomial model	$RSSI = 0.00005d^5 - 0.01d^4 - 0.81d^3 - 32.34d^2 - 638.87d - 4976$	35.2m

The log-normal model and the third-order polynomial model have insufficient fitting problems, while the fifth-order polynomial model has over-fitting problems. Therefore, the fourth-order polynomial model is selected for calculating the distance between nodes in subsequent node self-localization.

When the topology of the node is shown in Fig. 1, we use the selective anchor node method based on tetrahedral volume to obtain the relationship between the volume of the tetrahedron composed of the anchor node and the positioning error. As shown in Fig. 3, the volume of the tetrahedron composed of anchor nodes and the maximum value of the calculation error exist a significant negative correlation. Using the volume-based selective method can effectively reduce the positioning error. When the volume reach to 2.8m^3, the positioning error decrease to 2.3m.
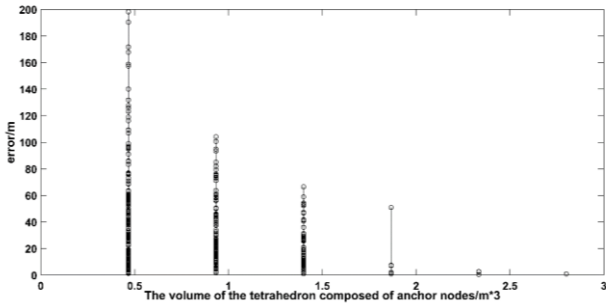
Fig. 3. The relationship between tetrahedral volume and positioning error

Finally, we compared the influence of whether to select anchor nodes on the node positioning algorithm in three-dimensional space. We selected 16 different nodes to be located and performed the different algorithms in turn, and the results obtained are shown in Figure 4. It can be seen from the figure that compared with randomly selecting anchor nodes, the method proposed in this paper improves the positioning accuracy by about 60%.
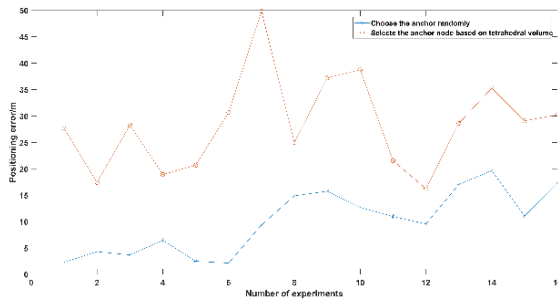
Figure. 4. Compare two different positioning method

So it can be obtained that the method proposed in this paper proved to be an effective DFL method which greatly improves the positioning accuracy of nodes in three-dimensional space.

V. CONCLUSION

This paper proposed a self-localization method that can select proper anchor nodes using the tetrahedron volume. Theoretical analysis and experiment verification have proved

a clear correlation between the volume of the tetrahedron and the positioning accuracy. Specifically, the larger the volume of the anchor node positioning unit, the higher the positioning accuracy. On the basis of this correlation, we propose a method of anchor nodes selection. The method of selecting anchor nodes based on the volume of tetrahedrons can realize the high-precision self-positioning of the nodes, which improves the efficiency of three-dimensional DFL system deployment.

REFERENCES

[1] Pirzada, Nasrullah , et al. "Device-free Localization Technique for Indoor Detection and Tracking of Human Body: A Survey." Procedia Social & Behavioral Sciences 129(2014):422-429.

[2] Mager B, Patwari N and Bocca M. Fall detection using rf sensor networks. In: Proceedings of the IEEE 24th international symposium on personal indoor and mobile radio communications (PIMRC), London, UK, 8‑11 September 2013, pp.3472‑3476. New York: IEEE.

[3] Anderson CR, Martin RK, Walker TO, et al. Radio tomography for roadside surveillance. IEEE J Sel Top Signa 2013; 8(1): 66‑ 79.

[4] Bocca M, Kaltiokallio O and Patwari N. Radio tomographic imaging for ambient assisted living. In: Chessa S and Knauth S (eds) Evaluating AAL systems through competitive benchmarking. Berlin: Springer, 2013, pp.108‑130.

[5] Patwari N, Brewer L, Tate Q, et al. Breathfinding: a wireless network that monitors and locates breathing in a home. IEEE J Sel Top Signa 2013; 8: 30‑42.

[6] Cui, Laizhong, Xu, Chong, Li, & Genghui. "A high accurate localization algorithm with DV-Hop and differential evolution for wireless sensor network." Applied Soft Computing (2018).

[7] Khalaf-Allah, Mohamed. "Time of Arrival (TOA)-Based Direct Location Method." International Radar Symposium IEEE, 2015.

[8] Bandiera, Francesco , A. Coluccia , and G. Ricci . "A Cognitive Algorithm for Received Signal Strength Based Localization." Signal Processing, IEEE Transactions on 63.7(2015):1726-1736.

[9] Kai, Zhao , L. I. Wei-Yao , and S. Ting . "RSS-Based Localization via Bayesian Ranging and Iterative Least Squares Positioning." Journal of Southwest China Normal University(Natural Science Edition) (2015).

[10] Kumar, Vikram, et al. "RSSI-Based Self-Localization with Perturbed Anchor Positions." IEEE International Symposium on Personal IEEE, 2017.

[11] Abouzar, Pooyan, D. G. Michelson, and M. Hamdi. "RSSI-Based Distributed Self-Localization for Wireless Sensor Networks Used in Precision Agriculture." IEEE Transactions on Wireless Communications (2016):1-1.

[12] Qiang, Rui, et al. "3D maximum likelihood estimation positioning algorithm based on RSSI ranging." 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) IEEE, 2017.

[13] Kirichek, Ruslan, et al. "Development of a node-positioning algorithm for wireless sensor networks in 3D space." 2016 18th International Conference on Advanced Communication Technology (ICACT) IEEE, 2016.

[14] Cheng S-W, Dey T K, Shewchuk J. Delaunay Mesh Generation [M], Chapman and Hall/CRC, 2012.

[15] Yan, Xiaoyong , H. Qian , and J. Yu . "A Selective Beacon Node 3D Location Estimation based on RSSI for Wireless Sensor Network." Information Technology Journal 12.1(2013):40-50.

Enhancing the Reliability of the Quadrotor by Formulating the Control System Model

Zhuo Zhi
Department of Test and Control Engineering
Harbin Institute of Technology
Harbin, China
zhuozhi@hit.edu.cn

Liansheng Liu
Department of Test and Control Engineering
Harbin Institute of Technology
Harbin, China
lianshengliu@hit.edu.cn

Datong Liu*
Department of Test and Control Engineering
Harbin Institute of Technology
Harbin, China
liudatong@hit.edu.cn

Abstract—**The quadrotor has been widely equipped in military and civilian areas. Due to small size and high integration, the system stability is easily affected by the environment. It is important to improve the reliability of the low-cost quadrotor control system and reduce the experimental risk. This paper proposes a method for formulating a simulation model of the flight control system and establishes a physical platform for verification. In this method, the Newton-Euler formula is involved to set up the kinematic equation. The Mahony filter is used for sensor data fusion and attitude solution. A cascade proportion integration differentiation controller with integral separation is designed to control the attitude. Finally, the experimental data are compared to the simulated data to verify the usability of this simulation model and the reliability of the flight control system.**

Keywords—reliability, quadrotor, simulation model, flight control system

I. INTRODUCTION

With the development of control theory, aero-dynamics theory, and sensor technology, the advantages of quadrotor have become prominent, such as simple structure, high flexibility, and low cost. It has been widely equipped for national defense and civil areas. In a quadrotor, data processing, command execution, and other tasks are all completed by the flight controller. The reliability of the flight controller determines the safety of the quadrotor to a great extent [1]. However, small sensors and large-scale integrated circuits would bring greater challenges because they are more sensitive to noise than their counterparts in larger vehicle. The reliability of the flight control system is gradually gaining attention [2-3].

It requires much cost and time to develop a flight control system for the quadrotor. Lots of experiments must be implemented to ensure robustness in unknown environments. The situation is compared to be more complex on a smaller scale, since it is more susceptible to temperature, wind speed, and magnetic strength. Besides, the vibration caused by motors can affect the data of micro electromechanical system (MEMS) sensors [4]. Therefore, formulating a simulation model to verify the reliability of the flight control system is helpful.

At present, most simulation models of the quadrotor flight control system are dedicated models, with strong pertinence, high complexity, and low degree of open source [5]. For example, semi-physical simulation is the most popular for its feasibility and practicality. However, it has less flexibility in the design of the control law and the adjustment of system parameters [6]. K. Derya et al. [7] use the wind tunnel to estimate the condition of hover, vertical climb, and forward flight modes. The simulation model is comprehensive. But, the cost is too high for general study. In summary, there is few research about the modeling method which is generally applicable to the low-cost quadrotor.

Regarding the problems aforementioned, this paper designs a simulation model of the flight control system and sets up a physical platform for experiments. The model includes three sub-models, the attitude control model, the motor dynamics model, and the rigid body dynamics model. The Newton-Euler formula is involved to set up the kinematic equation. The Mahony filter [7] is used for sensor data fusion and attitude solution. A cascade Proportion Integration Differentiation (PID) controller [8-9] with integral separation is designed to control the attitude. Finally, the reliability of the flight control system is proved by multi-mode control experiments on the quadrotor. The usability of this simulation model is verified through the comparison between the experimental data and simulation data.

This article is organized as follows. Section II introduces the methodology used in the design. Section III illustrates the components of the experimental setup. Section IV shows the experiments result under multi-control mode and compares the experimental data with simulation data. Section V draws the conclusion.

II. METHODOLOGY

This section introduces the methodology used in the design. The establishment of coordinate system and dynamic analysis is the basis to describe the motion of the quadrotor. The Mahony filter and the first-order complementary filter are used for attitude solution. The cascade PID controller is designed to control the attitude.

A. Coordinate System Definition and Conversion

Two coordinate systems are utilized to describe the motion of the quadrotor. The first is named as the space coordinate system, and the other is named as the body coordinate system. In the body coordinate system, the angle of rotation around the X axis is named as the PIT angle, the angle of rotation around the Y axis is named as the ROL angle, and the angle of rotation around the Z axis is called the YAW angle [10-11]. The pulling force generated by rotors can be directly described in the body coordinate system. However, gravity needs to be calculated in the space coordinate system. They can transfer to each other by the rotation matrix. The three-axis rotation matrix can be obtained by

$$\text{R}x(\phi) = \begin{bmatrix} cos\,\phi & sin\,\phi & 0 \\ -sin\,\phi & cos\,\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

---

*Datong Liu is the corresponding author. (e-mail: liudatong@hit.edu.cn).

$$Ry(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\,\theta & sin\,\theta \\ 0 & -sin\,\theta & cos\,\theta \end{bmatrix}, \qquad (2)$$

$$Rz(\psi) = \begin{bmatrix} cos\,\psi & 0 & -sin\,\psi \\ 0 & 1 & 0 \\ sin\,\psi & 0 & cos\,\psi \end{bmatrix}, \qquad (3)$$

where ϕ denotes the PIT angle, θ denotes the ROL angle, and ψ denotes the YAW angle.

If a vector P rotates around the *X*, *Y*, and *Z* axes in space, the relationship between the rotated vector P' and the previous vector is given in (4).

$$P' = Rx \cdot Ry \cdot Rz \cdot P \qquad (4)$$

The total rotation matrix characterized by the quaternion is shown in (5).

$$\begin{bmatrix} 1\text{-}2(q2q2\text{+}q3q3) & 2(q1q2\text{+}q0q3) & 2(q1q3\text{-}q0q2) \\ 2(q1q2\text{-}q0q3) & 1\text{-}2(q1q1\text{+}q3q3) & 2(q2q3\text{+}q0q1) \\ 2(q1q3\text{+}q0q2) & 2(q2q3\text{+}q0q1) & 1\text{-}2(q1q1\text{+}q2q2) \end{bmatrix} \quad (5)$$

The q0, q1, q2 and q3 are used to form a quaternion. The quaternion method is used to calculate the process of vector rotation, it can also avoid the problem of universal lock.

B. Establishment of the Kinematic Equation

The external forces and moments experienced by the quadrotor include three items.

- Gravity mg .

- Lift F_i (i=1, 2, 3, 4) generated by the rotation of rotors.

- Rotor rotation produces torsional M_i (i=1, 2, 3, 4), and it is perpendicular to the rotor plane of blades which is opposed to rotation vector.

The linear motion is caused by all the external forces, which accord with Newton's second law [12]. The linear motion equation is formulated in the space coordinate system, which is shown by

$$m\ddot{r} = \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix} + R_t \begin{bmatrix} 0 \\ 0 \\ \sum F_i \end{bmatrix}, \qquad (6)$$

where r is the position vector of the quadrotor, F_i is the lift generated by the rotation of rotors, mg is the gravity, R_t is the rotation matrix. And the angular motion equation is shown by

$$I \begin{bmatrix} \dot{p} \\ \dot{q} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} L(F_2\text{-}F_4) \\ L(F_3\text{-}F_1) \\ M_1 - M_2 + M_3 - M_4 \end{bmatrix} - \begin{bmatrix} p \\ q \\ r \end{bmatrix} \times I \begin{bmatrix} p \\ q \\ r \end{bmatrix}, \qquad (7)$$

where L is the distance from the center of the rotor to the center of mass of the quadrotor, I is the inertia matrix, p , q and r are components of triaxial angular velocity in the body coordinate system.

C. Attitude and Height Algorithm

The Mahony filter is utilized as the attitude algorithm. The core idea is based on predictive-correction. The main steps are as follows [13].

1) The standard gravity vector in the space coordinate system is converted into the body coordinate system by the rotation matrix. The process is shown as

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 2(q2q2\text{+}q3q3) & 2(q1q2\text{+}q0q3) & 2(q1q3\text{-}q0q2) \\ 2(q1q2\text{-}q0q3) & 1\text{-}2(q1q1\text{+}q3q3) & 2(q2q3\text{+}q0q1) \\ 2(q1q3\text{+}q0q2) & 2(q2q3\text{+}q0q1) & 1\text{-}2(q1q1\text{+}q2q2) \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} =$$
$$\begin{bmatrix} 2(q1q3\text{-}q0q2) \\ 2(q2q3\text{+}q0q1) \\ 1\text{-}2(q1q1\text{+}q2q2) \end{bmatrix}, \qquad (8)$$

where $[\,v_x \;\; v_y \;\; v_z\,]$ is the gravity vector after the standard gravity vector is converted to the body coordinate system. [0 0 1] is the coefficient of the standard gravity vector.

2) The error between the standard gravity vector converted in the previous step and the gravity vector measured in the body coordinate system is obtained by cross multiplication. The result is the error of the two coordinate systems. The specific step is shown by

$$\begin{bmatrix} e_x \\ e_y \\ e_z \end{bmatrix} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \otimes \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} a_y * v_z - a_z * v_y \\ a_z * v_x - a_x * v_z \\ a_x * v_y - a_z * v_x \end{bmatrix}, \qquad (9)$$

where $[\,a_x \;\; a_y \;\; a_z\,]$ is the gravity vector measured by sensors, and $[\,e_x \;\; e_y \;\; e_z\,]$ is the error vector of two coordinate systems.

3) The error vector obtained in the previous step is used to compensate for the gyroscope. A proportion integration controller is designed for this step.

The first-order complementary filter is used to calculate the speed, acceleration, and other parameters on the *Z* axis of the quadrotor. The detailed steps are explained as follows.

1) The acceleration a_j measured in the body coordinate system is converted to the space coordinate system to obtain the ground acceleration a_g .

2) The ground speed v is estimated by performing a first-order complementary fusion on the differential value of the barometer and a_g .

3) The height h is estimated by performing a first-order complementary fusion on the original data of the barometer and v .

The formula of the first-order complementary filtering is shown by

$$\theta_c(t) = \frac{\varepsilon}{\varepsilon + dt}(\theta_c(t-1) + \omega dt) + (1 - \frac{\varepsilon}{\varepsilon + dt})\theta(t), \quad (10)$$

243

where $\theta_c(t)$ is the current fusion result, $\theta_c(t-1)$ is the fusion result at the last moment, $\theta(t)$ is the data input at the current moment, ω is the change amount, and ε is the weight [14].

D. Attitude Control Method

The final actuators of the control system are four motors. The control of the attitude is achieved by adjusting the speed of motors. The PID controller is a kind of closed-loop controller based on the feedback deviation of output. This paper designs a double-loop cascade PID controller with integral separation. Take the ROL angle as an example, the block diagram is shown in Fig. 1.
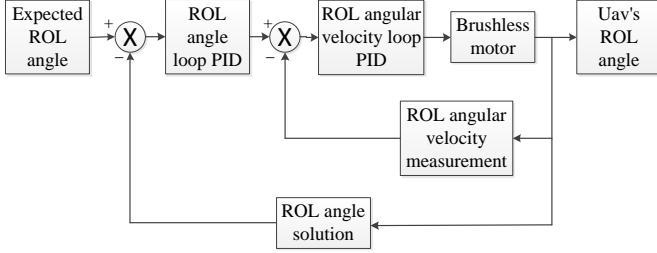
Fig. 1. Block diagram of cascade PID control for ROL angle

When the deviation is large, the integral term is cleared, as well it is turned on when the deviation is small, which can achieve rapid dynamic response and stable static control.

III. EXPERIMENT SETUP

The experimental setup includes the simulation model and the quadrotor. The main components of these setups are introduced in this section.

A. Components of the Simulation Model

The simulation model includes three sub-models, the attitude (including height and attitude) control model, the motor dynamics model, and the rigid body dynamics model. The structure diagram of the model is shown in Fig. 2.
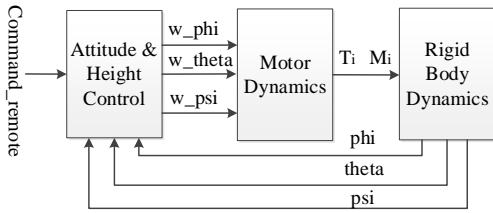
Fig. 2. Structure diagram of the simulation model

The input of the attitude & height control model is the desired attitude angle from the remote and the actual attitude angle, and the output is the desired speed of rotors. The input of the motor dynamics model is the desired speed of the rotor, and the output is the force and moment acting on the body. The input of the rigid body dynamics model is the pulling force and moment provided by rotors, and the output is the attitude angle of the quadrotor.

B. Components of the Quadrotor

The flight controller is designed according to the algorithm used in the simulation model and carried by a frame for experiments. The parameters of the frame are shown in TABLE I. The selection of the frame and power pack is based on the dynamic calculation in the simulation model.

TABLE I. PARAMETERS OF THE FRAME

No.	Index	Value
1	Wheelbase	330mm
2	Motor model	X2212-1400KV
3	Propeller diameter	20.3cm
4	Total weight	606g

The hardware design and the software design of the flight controller are shown in Fig. 3 and Fig. 4.
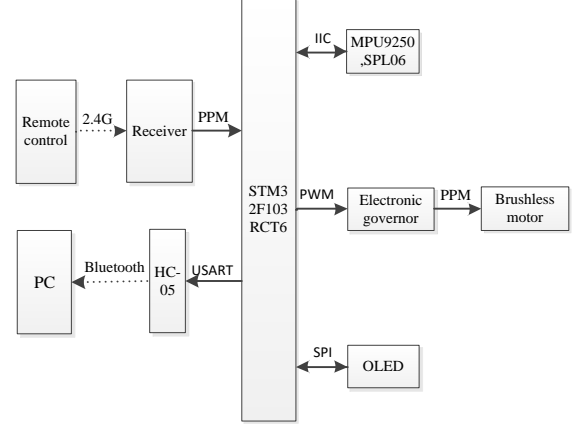
Fig. 3. Block diagram of hardware design

The model of the Microcontroller Unit (MCU) in this system is STM32F103RCT6. The used sensors are MPU9250 (with built-in gyroscope, accelerometer, and magnetometer), and SPL06 barometer.
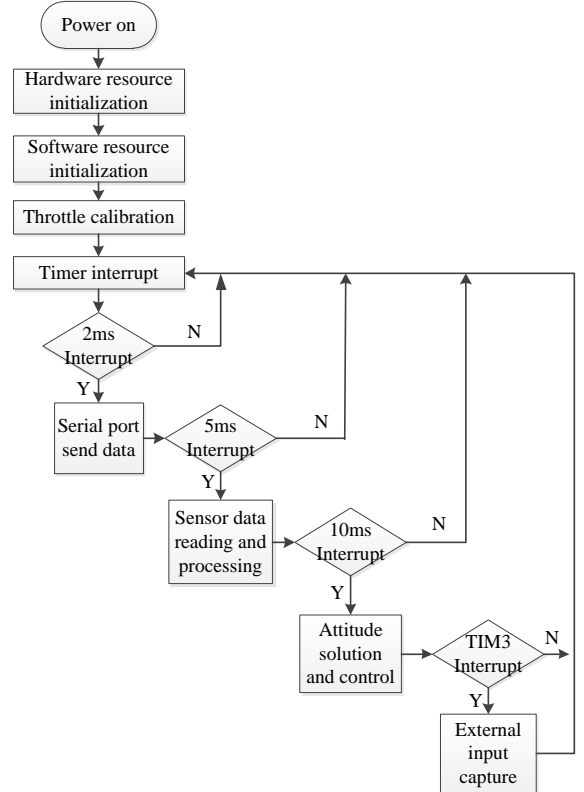
Fig. 4. Flow chart of system software

Due to limited chip resources, this design uses the idea of multi-thread. All software functions are divided into several tasks according to the frequency, such as upper computer communication, sensor data reading, processing, attitude

solution, attitude control, and external input capture. They are triggered by timers.

IV. EXPERIMENT AND COMPARATIVE ANALYSIS

The simulation model of the flight control system is verified in two aspects. The first is the height control and another is the altitude control. The attitude control includes the static and the dynamic attitude control. The same commands are input to the simulation model and the physical platform to get two groups of output data. Pearson correlation coefficient is calculated to analyze the correlation between two groups of data.

Assuming X and Y are two independent variables, the formula of the Pearson correlation coefficient is shown by

$$\rho_{X,Y} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{(X_i - \bar{X})^2(Y_i - \bar{Y})^2}}, \qquad (11)$$

where $\bar{X}$, $\bar{Y}$ are the mean of two groups of data, X_i, Y_i are the data elements in two groups of data, $\rho_{X,Y}$ is the correlation coefficient, which is between -1 and 1. The value of 1 indicates that the variables are completely positively correlated, 0 represents irrelevant, and -1 shows completely negatively correlated. A correlation coefficient greater than 0.8 indicates a strong positive correlation between two groups of data.

A. The Height Control Experiment

The main test objects of this experiment are the effect of the fixed height control and the variable height control. The height control instructions are as shown in TABLE II. The height data measured in the simulation model and the physical platform are shown in Fig. 5 and Fig. 6, respectively.

TABLE II.   HEIGHT CONTROL INSTRUCTIONS

No.	Time	Desired Height
1	3.0s	60cm
2	12.5s	35cm
3	18.0s	0cm
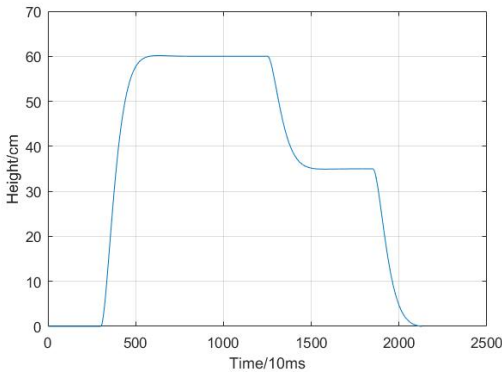
Fig. 5.   Height data measured in the simulation model

The curve of height data measured in the physical platform has the same trend as the curve of height data measured in the simulation model. The correlation between the two sets of data is 0.9794 calculated by the Pearson correlation coefficient. It can be regarded that the output of the simulation model and physical platform in the height control experiment is consistent.
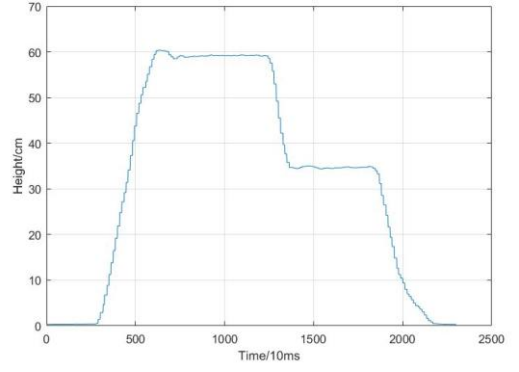
Fig. 6.   Height data measured in the physical platform

B. The Static Attitude Control Experiment

The main test objects of the static attitude control experiment are static error and floating range of the attitude angle when the quadrotor is hovering. The angle data measured in the simulation model and the physical platform are shown in Fig. 7 and Fig. 8, respectively.
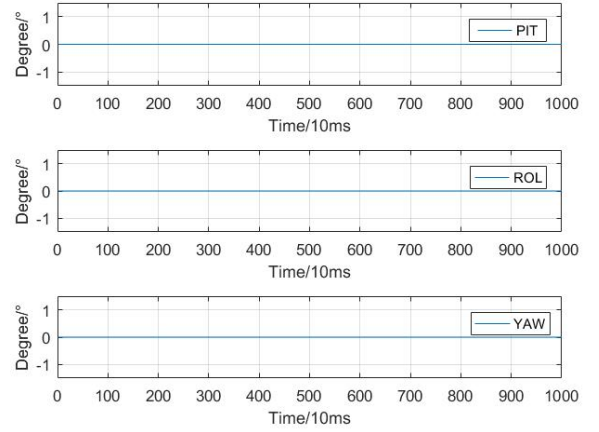
Fig. 7.   Angle data measured in the simulation model
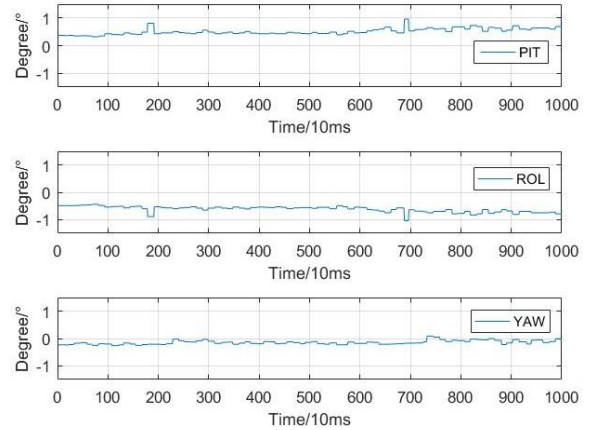
Fig. 8.   Angle data measured in the physical platform

The duration of the hovering experiment is 10s. The attitude angle output of the simulation model is all 0. Due to mechanical vibration, internal noise, and other factors in the physical platform, the attitude angle fluctuates slightly. But the range does not exceed 1 degree. Because of the incomplete level of the take-off plane, the attitude angle output of the physical platform has a certain static error, and it remains stable. The variance of three sets of data of the physical

platform are 0.0124, 0.0135, and 0.0058, respectively. To sum up, the output of the simulation model and physical platform in the static attitude control experiment is consistent.

C. The Dynamic Attitude Control Experiment

The main test objects of the dynamic attitude control experiment are the overshoot, the response speed, and the regression error. During the test, the desired angle input is set randomly. The angle data measured in the simulation model and the physical platform are shown in Fig. 9 and Fig. 10, respectively.
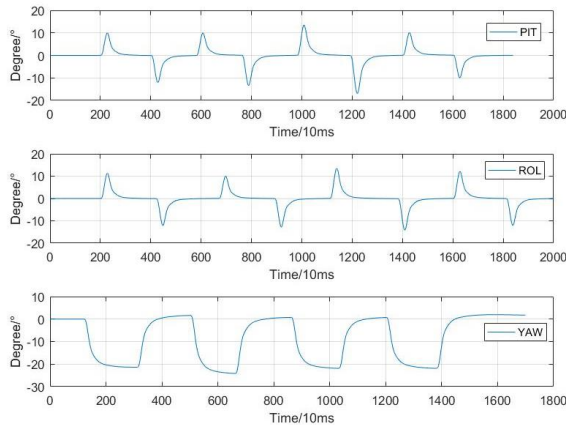
Fig. 9. Angle data measured in the simulation model
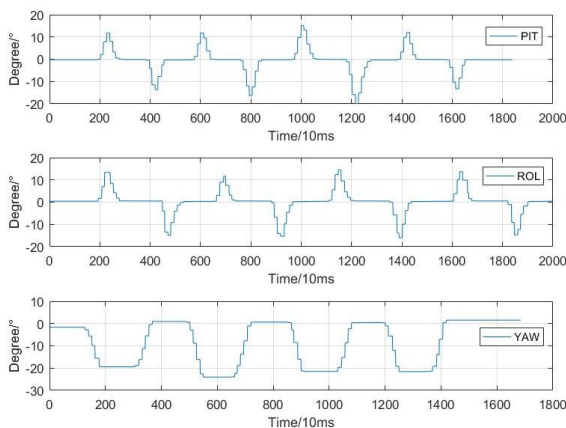
Fig. 10. Angle data measured in the physical platform

The experimental waveforms can well follow the simulated waveforms, and no runaway or oscillation occurs. Due to sensor noise and unavoidable environmental factors, there are small glitches in the measured data. The control system is accompanied with overshoot, which does not exceed 2 degrees. The regression error does not exceed 1 degree after the end of each instruction cycle. And the delay of controller response is less than 10ms. The correlation of three sets of comparative data are 0.9098, 0.9192, and 0.9688, respectively. Therefore, it can be considered that the output of the simulation model and physical platform in the dynamic response attitude control experiment is consistent.

V. CONCLUSION

This paper proposes a method for formulating a simulation model for the low-cost quadrotor flight control system. A physical platform is set up for experiments. The experimental results show that the method can accurately describe the actual motion state of the quadrotor and verify that the flight control system has strong stability. Hence, this method can be applied to the research and development of the quadrotor flight control system, which can effectively improve the reliability of the flight control system design.

REFERENCES

[1] M. Shamim Hossain, F. Alhamid, "MT-AAAU: Design of Monitoring and Tracking for Anti-Abuse of Amateur UAV," Mobile Networks and Applications, vol. 23, no. 2, pp. 328–335, 2018.

[2] H. Mo, G. Farid, "Nonlinear and Adaptive Intelligent Control Techniques for Quadrotor UAV–A Survey," Asian Journal of Control, vol. 21, no. 2, pp. 989–1008, 2019.

[3] L. Liu, M. Liu, Q. Guo, D. Liu, Y. Peng, "MEMS Sensor data anomaly detection for the UAV flight control subsystem," IEEE SENSORS Conference, 2018, pp. 1–4.

[4] K. Büyükkabasakal, B. Fidan, A. Savran, "Mxing Adaptive Fault Tolerant Control of Quadrotor UAV," Asian Journal of Control, vol. 19, no. 4, pp. 1441–1454, 2018.

[5] Q. Ge, T. Shao, Z. Duan, C. Wen, "Performance Analysis of the Kalman Filter With Mismatched Noise Covariances," IEEE Transactions on Automatic Control, vol. 61, no. 12, pp. 4014–4019, 2016.

[6] A. Miller, "Developing algorithms of object motion control on the basis of Kalman filtering of bearing-only measurements," Autom Remote Control, vol. 76, no. 6, pp. 1018–1035, 2015.

[7] K. Derya, A. Kutay, "Aerodynamic Modeling and Parameter Estimation of a Quadrotor Helicopter," AIAA Atmospheric Flight Mechanics Conference, 2014, pp. 2514–2558.

[8] A. Moutinho, J.R. Azinheira, "Attitude Estimation in SO(3): A Comparative UAV Case Study," Journal of Intelligent & Robotic Systems, vol. 80, no. 3-4, pp. 375–384, 2015.

[9] K. J. Aström, T. Hägglund, "Nichols step response method for PID control," Journal of Process Control, vol. 14, no. 6, pp. 635–650, 2004.

[10] H. Hjalmarsson, M. Gevers, "For modelbased control design, closed-loop identification gives better performance," Automatica, vol. 32, no. 12, pp. 1659–1673, 1996.

[11] W. Chao, X. Yi, "Tuning of sampled-data ADRC for nonlinear uncertain systems," Journal of Systems Science and Complexity, vol. 29, no. 5, pp. 1187–1211, 2016.

[12] W. Xue, Y. Huang, "On frequency-domain analysis of ADRC for uncertain system," Proceedings of the 2013 American Control Conference, 2013, pp. 425–455.

[13] O. N. Bogdanov, A. V. Fomichev, "Effect of Delays in Rate Sensor Channels on the Accuracy of Navigation Solution of Strapdown Inertial Navigation System," Gyroscopy and Navigation, vol. 9, no. 3, pp. 199–218, 2018.

[14] A. Sanchez, L. R. García Carrillo, "Hovering Flight Improvement of a Quad-rotor Mini UAV Using Brushless DC Motors," Journal of Intelligent & Robotic Systems, vol. 61, no. 1-4, pp. 85-101, 2011.

[15] Z. Shiyu, L. Feng, "Vision-aided Estimation of Attitude, Velocity, and Inertial Measurement Bias for UAV Stabilization," Journal of Intelligent & Robotic Systems, vol. 81, no. 3–4, pp. 531–549, 2016.

[16] E. I. Pounds, R. Bersak, M. Dollar, "Stability of small-scale UAV helicopters and quadrotors with added payload mass under PID control," Autonomous Robots, vol. 33, no. 1–2, pp. 129–142, 2012.

[17] D. Zhi, B. Jiang, P. Shi, "Global robust tracking control of non-affine nonlinear systems with application to yaw control of UAV helicopter," International Journal of Control, Automation and Systems, vol. 11, no. 5, pp. 957–965, 2013.

[18] Q. Jun, Han. Jin, "Application of wavelets transform to fault detection in rotorcraft UAV sensor failure," Journal of Bionic Engineering, vol. 4, no. 4, pp. 265–275, 2007.

[19] K. K. Veremeenko, V. M. Savel'ev, "In-flight alignment of a strapdown inertial navigation system of an unmanned aerial vehicle," Journal of Computer and Systems Sciences International, vol. 52, no. 1, pp. 106–116, 2013.

An SOC and SOP Joint Estimation Method of Lithium-ion Batteries in Unmanned Aerial Vehicles

Wanqing Cheng
Automatic Test and Control Institute
Harbin Institute of Technology
Harbin, China
chengwanqing88@163.com

Zhiheng Yi
School of Instrument Science and
Engineering
Harbin Institute of Technology
Harbin, China
984927321@qq.com

Jun Liang*
School of Electronics and Information
Engineering
Harbin Institute of Technology
Harbin, China
hitliangjun2012@126.com

Yuchen Song
School of Electronics and Information Engineering
Harbin Institute of Technology
Harbin, China
songyuchen@hit.edu.cn

Datong Liu
School of Electronics and Information Engineering
Harbin Institute of Technology
Harbin, China
liudatong@hit.edu.cn

Abstract—**Unmanned aerial vehicles (UAVs) are widely used in industry and military scenarios due to the advantages of low-cost. For drones using lithium-ion batteries as the energy storage component, accurate state estimation plays an important role in the properly decision making and safely control and maintenance. However, under dynamic operating conditions, only one state is often estimated, in fact there is often a certain correlation between each state, ignoring this will lead to inaccurate estimation results. To address these challenging issues, this paper proposes a joint estimation method for lithium-ion battery's state-of-charge (SOC) and state-of-power (SOP). Firstly, a method based on Long Short-Term Memory (LSTM) is used to predict the SOC of lithium batteries. There is no need to build a complex battery model. Based on the prediction results, according to the coupling relationship between SOC and SOP, the multi-parameter constraint method is used to obtain the estimated SOP value in a period of time. The experiment was conducted based on the battery simulation data. The results based on the lithium-ion battery simulation data show that the method used in this paper can accurately predict the battery SOC and SOP state parameters.**

Keywords—Unmanned aerial vehicles (UAVs), Lithium-ion battery, Long-short term memory (LSTM), State of Charge (SOC), State of Power (SOP), Joint state estimation

I. Introduction

As an efficient and convenient auxiliary method, UAV has been widely used in the fields of electric power inspection, emergency rescue, aerial photography, pesticide spraying, etc. It has advantages in low cost, good cost-effective ratio, good mobility, etc. which reduces the risk of manual operations and improves the maneuverability of task execution [1]. At the same time, taking the advantages of high power density, lithium-ion batteries are also widely used to supply stable electric power. Especially for those electric UAVs (i.e. quadcopter, hexacopter, etc.), lithium-ion batteries are the only power source. Therefore, the states of them directly influence the reliability and operation strategy of the UAVs.

Aiming at different flying missions of UAVs, it is meaningful to accurately evaluate the ability of completing a specific task and design the flight strategy [2]. For lithium-ion batteries, two of the most important state parameters are SOC and SOP [3]. Accurate estimation of these two parameters can ensure safe flight of the UAV and helps users to make proper flight decision as well as effective maintenance actions [4].

SOC is defined as the ratio of the current capacity and the available capacity of the battery [5]. During the working process of the UAV, when the battery SOC is insufficient, there may be problems such as interruption of communication and limited operating capacity [6].The measurement methods for SOC mainly include Coulomb counting method, open-circuit voltage (OCV) method, load voltage method, internal resistance method and the Kalman filter method [7,8]. Recently, machine learning method can also be applied to battery SOC estimation. This method does not require the establishment of complex lithium-ion battery models. Based on the historical samples of external parameters and corresponding SOC, the estimator is trained and applied to estimate the current state [8]. Methods such as support vector machines(SVM) and neural networks all belong to this category. Its advantage is that it does not rely too much on battery's model, only a certain amount of training data is needed, and it can approximate any non-linear model. And the parameters in the data-driven model are automatically identified in practical operating conditions. The extra testing procedures are no longer needed for data-driven SOC estimation methods [5].

SOP represents the ratio of peak power to nominal power. Under the voltage, current, SOC, and power limitations, the maximum power that a battery can persistently provide for a specific time duration is defined as the peak power [3]. Sufficient power should be provided in the stage of climb acceleration [5]. At present, the estimation method for SOP is mainly divided into battery experiment method and parameter constraint method [5]. The experimental method is HPPC method. It calculates the peak power based on the open circuit voltage and the internal resistance [9]. The testing process of it is simple to implement. But the battery is not discharged to the cut-off voltage, so the power obtained is less than the actual value. Parameter constraint method mainly includes voltage constraint method and SOC constraint method. United States Partnership for a New Generation of Vehicle (PNGV) composite pulse method is a method based on voltage constraints. When estimating the SOP, the peak current of charge and discharge is calculated with the charge and discharge cut-off voltage in the battery as the limiting condition. The designed or rated battery current and SOC is not considered [7]. The SOC-limited estimation method mainly considers the influence of battery SOC limitation in power calculation. Therefore, the influence is greater only when the battery reaches the cut-off SOC. Therefore, this method is often combined with other methods. The main method of calculation is to obtain the battery peak current based on the battery discharge current, voltage, SOC and other limitations, thereby obtaining the battery peak power [10].

For the estimation of SOC, some references proposed machine learning and Kalman filter (KF) methods to estimate SOC of the lithium-ion battery under dynamic conditions [9,11]. but only the SOC can be estimated in these frameworks.. In practical applications, it cannot provide comprehensive basis for decision-making. References [5] realize the joint estimation of battery SOC and SOP, but the research object is electric vehicles, and its working conditions are not applicable to drones. The difference is that the UAV's power system does not contain a power feedback system. Therefore, the battery will never be charged during the whole flight process. References [2, 9, 12] focused on the estimation of battery SOP, but in SOP estimation, the peak power under SOC limit requires accurate SOC estimation value, and the accuracy of SOC affects the SOP estimation result.

Regarding the issues above, how to realize the joint estimation of the SOC and SOP of the UAV's lithium-ion battery under its unique operating conditions is still a challenging issue. This paper implements a joint estimation of battery state based on long short-term memory (LSTM) method and multi-parameter constraint method. The LSTM algorithm does not require the establishment of a complicated battery model to make the result more accurate. At the same time, the multi-parameter constraint method is used to estimate the battery SOP, which can obtain more accurate results than one-parameter constraint. Both SOC and SOP estimation results can provide a reference for the UAV trajectory planning and can further enhance the operating safety and reliability of UAVs. The main contribution of this work is listed as follows:

- First, a machine learning method called LSTM network is proposed to estimate battery SOC, which solves the problem of building complex battery models.

- Second, a multiple-constraints method is used to estimate battery SOP, which can provide an effective reference for drone flight strategy

- Third, in this paper, joint state estimation of SOC and SOP is realized. According to the coupling relationship between two parameters, the result of SOC estimation can be used as a parameter of SOP estimation.

The rest of this paper is organized as follows. The battery simulation model and the flight simulation is given in Sect. II. And the joint state estimation of SOC and SOP is given in Sect. Ⅲ. Moreover, experiment and verification are showed in Sect. Ⅳ. In the end, conclusions are given in Sect. Ⅴ.

II. Data Set And Simulation Environment

A high precision lithium-ion battery electrochemical model is used to simulate the battery discharge behavior during the UAV flying. The initial SOC of the battery is set to nearly the highest value, 85.51%. Subsequent test data are all obtained using this method.

The working conditions selected in the experiment consists of the typical UAV flying process such as taking-off, cruise, hover, and landing. According to the operating characteristics of the UAV lithium-ion battery, the discharge method of segmented constant current discharge is set. First, targeting the drone climbing phase the current is discharged at 50A for 300s, and then during level flight, it discharges at 30A to 50% SOC. Stop discharging for 180s to simulate task execution, and then discharge at 50A to 30% SOC, finally apply 20A to discharge to 10% which reaches the cut-off condition during landing stage [13]. The changes in battery current, SOC, and voltage during charging and discharging under this condition are shown in Fig. 1 (a)-(c).

Fig. 1. Current simulation results. (a) Current simulation results, (b) SOC simulation results, (c) Voltage simulation results

III. Joint State Estimation

This chapter introduces the methods of SOC and SOP joint estimation. The SOC estimation model is developed based on the LSTM algorithm which uses voltage and current as input data. And the SOC is output. Next is SOP estimation, which requires calculation of open circuit voltage, internal resistance, etc., all obtained through simulation experiments. When calculating the peak current with SOC as the limiting condition, the estimation result of SOC needs to be used.

A. State-of-Charge Estimation

In this part, a data-driven method which consists of LSTM and Kalman filter algorithm is proposed to predict SOC. As regard to the choice of characteristic parameters, we should consider parameters like voltage or current which can be acquired more easily than those complex parameters like

resistance, which needs to be measured with accurate model or static measurement [15]. In this work, voltage and current are used as input parameters and the output is SOC. The framework of the model is shown in Fig. 2.
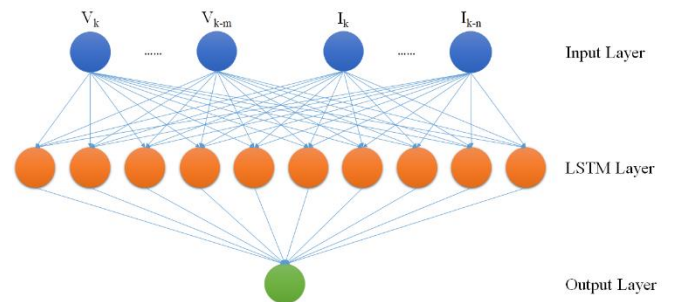
Fig. 2. The structure of the LSTM network model

LSTM network is good at time series prediction, because the forget gate in this network provide it with the ability to forget some features far away from the current time [16]. So we chose it as the network model for SOC prediction [13]. The results of prediction can be affected due to measurement noise or the sudden change in current value, so we apply Kalman filter to optimize the results [17]. The output value is used as the observation value of the state-space model, which can avoid the establishment of complex observation equations. Specific steps are as follows.

1) The reconstruction of the input data. We normalize the input data V, I and output data SOC within the range of [0, 1] to avoid the influence of large outliers on the model. The array V and I are reconstructed into $[v_k,...,v_{k-m},i_k,...,i_{k-n}]$ as the input vector when training the model. The embedding dimension m is selected as 30, and the embedding dimension n is selected as 10. This selection can avoid the influence of some special points in the input data on the model.

2) Batch-size and epochs is also set to 100. We choose the mean square error (MSE) as loss function, and at the time of 100 epochs, the loss has converged to a satisfied level. ReLU is used as activation function to make the model converge quickly. This activation function also solve the problem of vanishing gradient. Adam is chosen as the optimizer to minimize the total loss of the model. After the training process, save parameters of the model for future use.

3) Predict with the trained model. The input of this model is still expressed as $[v_k,...,v_{k-m},i_k,...,i_{k-n}]$.

4) Carry out Kalman filtering on the prediction results.

B. State-of-Power Estimation

To realize the SOP estimation, first we need to establish an electrochemical model and obtain relevant parameters. We choose Thevenin model, which considers the influence of battery polarization effect and can reflect the dynamic characteristics of the battery. The circuit is shown in Fig. 3.
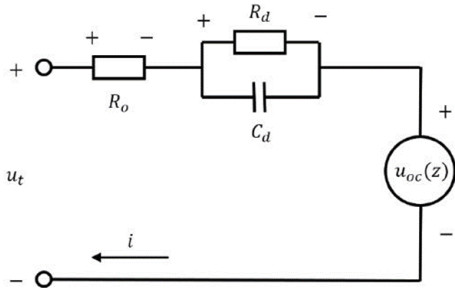
Fig. 3. The structure of the LSTM network model

The establishment of the model is [16]

$$\dot{U}_p = -\frac{1}{C_p R_p} U_p + \frac{1}{C_p} I_L \quad (1)$$

$$U_t = U_{oc} - U_p - I_L R_0 \quad (2)$$

where U_p is polarization voltage, C_p and R_p are the polarization capacitance and polarization resistance of the battery, U_{oc} is open circuit voltage, which is related to SOC, I_L is the battery current, R_0 is internal resistance.

In the model, U_{oc}, I_L and R_0 are all measured by experiment, C_p and R_p are all unknown parameters. All need to be obtained by the method of discrete parameter identification.

Next is an introduction to solving model parameters.

1) *Open Circuit Voltage:* It is obtained in advance through offline experiments. This work uses the interval standing method. First, charge the battery to a fully charged state with constant current and constant voltage, and let it stand for a period of time. Using a current density of 1C to discharge the battery at a constant current. Every time the battery discharges 10% SOC with standing for 30 minutes. Then the terminal voltage of the battery is the open circuit voltage of the battery under this SOC [3]. Assuming that the relationship between OCV and SOC is [5, 6].

$$U_{OCV} = k_0 + k_1 SOC + k_2 / SOC + k_3 \ln SOC + k_4 \ln(1-SOC) \quad (3)$$

The $\partial U_{OCV} / \partial SOC$ can be obtained from Eq.3. And the identified parameters are brought into the resulting expression, and the result is [16]

$$\frac{\partial U_{OCV}}{\partial SOC} = k_1 - k_2 / SOC^2 + k_3 / SOC - k_4 / (1-SOC) \quad (4)$$

2) *Internal Resistance:* The measurement of the internal resistance of the battery is designed according to the HPPC test method in the American Freedom CAR Battery Experiment Manual.

3) *Polarization Parameters:* They mainly include battery polarization capacitance and polarization resistance. In the battery polarization model, with current as input, the battery terminal voltage is [7]

$$U_t = U_{oc} + R_0 I + R_p I(1-\exp(-t / R_p C_p)) \quad (5)$$

Choose exponential form for fitting, the terminal voltage expression is

$$\hat{U}_t = k_0 + k_1 e^{-\lambda t} \quad (6)$$

where k_0, k_1, λ are parameters to be fitted. We use least squares fitting parameters. Then C_p and R_p are as follows

$$R_p = \frac{k_1}{I} \quad (7)$$

$$C_p = \frac{1}{\lambda R_p} \quad (8)$$

After solving the parameters. Then using SOC, voltage and current as the limiting conditions to predict the SOP of the battery in the next 2min under different SOC. The following is the multi-parameter constraint method to estimate SOP.

1) *SOC Constraint:* In order to prevent the battery from overcharging and discharging, the upper and lower constraints of the battery are usually limited to cut off the SOC. During

battery operation, we should make the battery work in a safe range. According to the Coulomb counting method, the SOC of the battery is [9].

$$s(t + \Delta t) = s(t) - I_L(t)(\frac{\eta_i \Delta t}{C_{max}}) \qquad (9)$$

where $s(t)$ is the battery SOC at time t. I is the battery current, η_i is the effective coefficient, which is the ratio of the discharge capacity to the charge capacity. C_{max} is the current maximum battery capacity.

Assuming that the minimum SOC of the battery is s_{min} and the maximum SOC is s_{max}, the current limit of the battery is [19]

$$I_{min}^{chg,soc} = \frac{s(t) - s_{max}}{\eta_i \Delta t / C_{max}} \qquad (10)$$

$$I_{max}^{dis,soc} = \frac{s(t) - s_{min}}{\eta_i \Delta t / C_{max}} \qquad (11)$$

The disadvantage of this method is that when the battery SOC range is large, it does not have a reference value. Therefore, this method is usually used when the current battery SOC value is close to the critical state, and is used together with other methods.

2) *HPPC Method:* The method was developed by the Idaho National Engineering and Environmental Laboratory and based on the Rint model [18].

$$U_t(t) = U_{oc}(s(t)) - R_0 I_L(t) \qquad (12)$$

Where $U_t(t)$ is the battery terminal voltage at time t, $U_{oc}(s(t))$ is the open circuit voltage at the current SOC, R_0 is the internal resistance of the battery, and $I_L(t)$ is the battery current at t. According to the battery voltage limit $U_{t,min} \leq U_t(t) \leq U_{t,max}$, the peak current of the battery under charging and discharging is

$$I_{min,k}^{chg,HPPC} = \frac{U_{oc,k} - U_{t,max}}{R_0} \qquad (13)$$

$$I_{max,k}^{dis,HPPC} = \frac{U_{oc,k} - U_{t,min}}{R_0} \qquad (14)$$

So the peak power is

$$P_{min}^{chg} = U_{t,max} \times I_{min}^{chg,HPPC} \qquad (15)$$

$$P_{max}^{dis} = U_{t,min} \times I_{max}^{dis,HPPC} \qquad (16)$$

This method is suitable for the laboratory, and its model is too simple, so the measured value is very different from the real value, so it is not suitable for using in actual situations.

3) *Voltage Constraint:* When the battery is at a constant voltage limit, according to the battery equivalent circuit model,

it is necessary to predict the peak charge and discharge current of the battery under the given limit conditions. According to the battery equivalent circuit model, the terminal voltage of the battery at time $t + \Delta t$ can be obtained as follows

$$U_t = U_{oc} - U_p - i_L R_0 \qquad (17)$$

Discretize the above formula is

$$U_p(t + \Delta t) = \exp(-\Delta t / \tau)U_p(t) + R_p(1 - \exp(-\Delta t / \tau))i_L(t) \qquad (18)$$

Then the peak current is [7]

$$i_{max}^{dis,volt} = \frac{U_{oc}(s(t)) - U_p(t)\exp(-\Delta t / \tau) - U_{t,min}}{\frac{\eta_i \Delta t}{C_{max}} \frac{\partial U_{oc}(s)}{\partial s}\Big|_{s=s(t)} + R_p[1 - \exp(-\Delta t / \tau)] + R_{dis}} \qquad (19)$$

$$i_{min}^{chg,volt} = \frac{U_{oc}(s_k(t)) - U_{p,k}(t)\exp(-\Delta t / \tau) - U_{t,max}}{\frac{\eta_i \Delta t}{C_{max}} \frac{\partial U_{oc}(s)}{\partial s}\Big|_{s=s(t)} + R_p[1 - \exp(-\Delta t / \tau)] + R_{chg}} \qquad (20)$$

$i_{max}^{dis,volt}$ and $i_{min}^{chg,volt}$ are the maximum discharge current and the minimum charge current of the battery, respectively. The charge current is specified as positive, and the discharge current is negative. τ is the battery time constant, R_{chg} and R_{dis} are the battery charge and discharge resistance.

Combining the above constraints, the peak current are [7]

$$i_{max}^{dis} = \min(i_{max}, \min i_{max}^{dis,SOC}, \min i_{max}^{dis,volt}) \qquad (21)$$

$$i_{min}^{chg} = \max(i_{max}, \max i_{min}^{chg,SOC}, \max i_{min}^{chg,volt}) \qquad (22)$$

According to the battery model, the corresponding terminal voltage is obtained, and finally the peak power of the battery in a period of time is [5]

$$P_{max}^{dis} = \sum_{k=1}^{n_s} i_{max}^{dis} U_{t,k}(t + \Delta t)$$
$$= (U_{OC}(Z_k) - \exp(-\frac{\Delta t}{\tau})U_{D,k} - I_{max}^{dis}(\frac{\Delta t \eta_i}{C_{max}} \frac{\partial U_{OC,k}}{\partial z}\Big|_{Z=Z(k)} \qquad (23)$$
$$+ R_{dis} + (1 - \exp(-\frac{\Delta t}{\tau}))R_D))I_{max}^{dis}$$

$$P_{min}^{chg} = \sum_{k=1}^{n_s} i_{min}^{chg} U_{t,k}(t + \Delta t)$$
$$= (U_{OC}(Z_k) - \exp(-\frac{\Delta t}{\tau})U_{D,k} - I_{min}^{chg}(\frac{\Delta t \eta_i}{C_{max}} \frac{\partial U_{OC,k}}{\partial z}\Big|_{Z=Z(k)} \qquad (24)$$
$$+ R_{dis} + (1 - \exp(-\frac{\Delta t}{\tau}))R_D))I_{min}^{chg}$$

C. Joint State Estimation

According to the above description, we can conclude the method of joint state estimation, which is shown in Fig.4. First is the estimation of SOC, which is the yellow part of the figure. The input data is the current and voltage data, and the output data is SOC. Then the result is subjected to Kalman filtering to obtain the final estimated values. The green part is the relevant parameter of the battery equivalent circuit model to

be solved. Using the obtained parameters and SOC estimation results to realize the peak current estimation of battery multi-parameter constraints with SOC constraint in (10) and (11), and voltage constraint in (19) and (20).And considering about the limit of current, SOP estimation can be solved. As a result, we can realize the joint estimation of SOC and SOP.
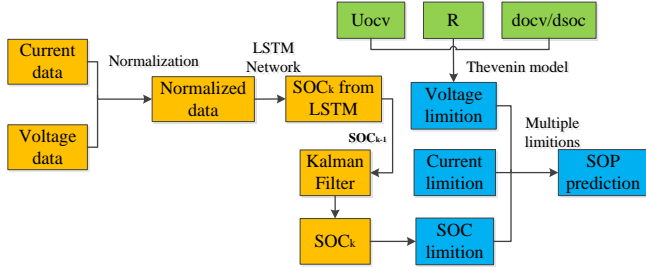
Fig. 4. The structure of the joint state estimation method

When estimating the SOP of the battery, the limitation of the battery needs to be provided. The following table lists the battery voltage, current, and SOC limits.

TABLE I. BATTERY'S RESTRICTIONS

extremum	Limited conditions		
	U/V	I/A	SOC/%
maximum	4.17	60.00	88.00
minimum	2.50	-30.00	10.00

IV. EXPERIMENT AND VERIFICATION

This chapter realizes the solution of battery parameters, the results comparison and error analysis of multi-dimensional state joint estimation.

A. Parameters' Measurement and Identification

Open circuit voltage results obtained by establishing equations and parameter identification are shown in Fig. 5.
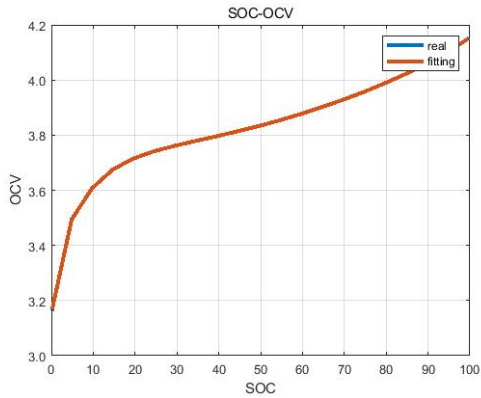
Fig. 5. OCV in function of SoC for the battery module.

The values of the parameters obtained by curve fitting are shown in the TABLE II.

TABLE II. VALUE OF PARAMETERS

extremum	parameters				
	k_0	k_1	k_2	k_3	k_4
value	3.491	0.002	-0.001	0.090	-0.026

The $\partial U_{OCV} / \partial SOC$ result is shown in Fig. 6.
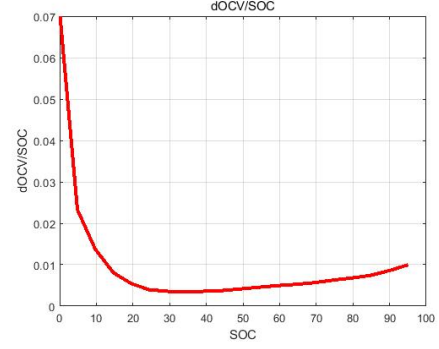
Fig. 6. Derivative of OCV as a function of SoC

The calculation result of resistance is shown in Fig. 7.
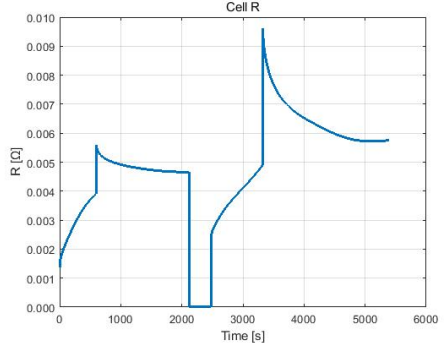
Fig. 7. Curve of battery internal resistance with time

B. Estimated Results

Use LSTM network training to get the prediction results and Error curve are shown in the Fig. 8.
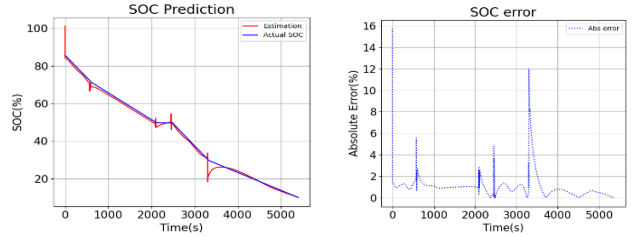
Fig. 8. SOC prediction. (a) Using LSTM model. (b) The error of SOC prediction

For the SOP verification method, a group of batteries of the same type and aging degree are used to test the constant power discharge capacity of the battery SOC at 10%, 15%, 20%, ... , 85%. The estimation results of the two methods are shown in Fig. 9.
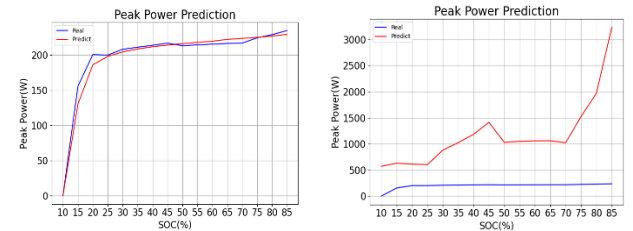
Fig. 9. PeakPower prediction. (a) Multiple limition method. (b) HPPC method

The evaluation indicators for the estimation results are RMSE and MAE. The results are shown in the table below. We can conclude that LSTM with KF method and multi-Restraint method can estimate SOC and SOP more accurately.

TABLE III.   EVALUATION OF ESTIMATED RESULTS

Different methods	Error	
	RMSE	*MAE*
SOC(without KF)	1.272	1.185
SOC(with KF)	8.055	1.038
SOP(HPPC method)	1160.503	9.8261
SOP(Multi-Restraint)	8.056	5.262

V.   CONCLUSIONS

This paper proposes a joint estimation method of multi-dimensional state of the UAV energy system under the dynamic conditions. SOC estimation is based on the LSTM network structure, with the battery voltage and current as input feature data, training to obtain SOC output. The data are obtained through electrochemical simulation experiments. Then using the coupling relationship between SOC and SOP to realize SOP estimation under multi-parameter constraints. The main constraints are battery SOC, voltage, and current. Finally, the estimation results are verified.

From the results and error analysis, it can be seen that the LSTM method with KF can be used to estimate SOC more accurately, which provides an important parameter value for the battery SOP estimation, making SOP estimation more accurate. And for SOP estimation, multi-constraint estimation is used. Compared with HPPC method, this method is more accurate. However, this method can not consider the degradation of battery life. So if we want to predict states in many battery circles, it would not be accurate enough. In the future, we can take the impact of SOH into account.

REFERENCES

[1] B. Wang, D. Liu, Y. Peng, and X. Peng, "Multivariate Regression-Based Fault Detection and Recovery of UAV Flight Data," IEEE Transactions on Instrumentation and Measurement, vol. 69, pp. 3527-3537, June 2020.

[2] Y. Song, D. Liu, H. Liao, and Y. Peng, "A hybrid statistical data-driven method for on-line joint state estimation of lithium-ion batteries". Applied Energy, 2020, vol. 261, pp. 114408.

[3] J. Jiang, S. Liu, Z. Ma, L. Wang, and K. Wu, Butler-Volmer equation-based model and its implementation on state of power prediction of high-power lithium titanate batteries considering temperature effects, Energy, 2016, vol. 117, pp. 10.1016.

[4] Y. Zhang, L. Wang, T. Wu, and M. Zhang, Unmanned aerial vehicle SOC estimation analysis based on improved EKF.Power Technology, 2019, vol. 43(02), pp. 320-323.

[5] R. Xiong, H. He. F. Sun, X Liu, and Z. Liu, Model-based state of charge and peak power capability joint estimation of lithium-ion battery in plug-in hybrid electric vehicles. Journal of Power Sources, 2013, vol. 229, pp. 159–169.

[6] R. Xiong, F. Sun, H. He, Nguyen, Trong Duy, A data-driven based adaptive state of charge estimator of lithium-ion polymer battery used in electric vehicles. Energy. 2013, vol. 63. pp. 295–308.

[7] Danko, Matúš, Adamec, Juraj, Taraba, Michal, Drgona, and Peter, Overview of batteries State of Charge estimation methods. Transportation Research Procedia. 2019, Vol. 40, pp. 186-192.

[8] X. Hu, F. Sun, and Y. Zou. Comparison between two model-based algorithms for Li-ion battery SOC estimation in electric vehicles. Simul. Model. Pract. Theory. vol. 34, pp. 1-11, 2013.

[9] Plett, Gregory L., "High-performance battery-pack power estimation using a dynamic cell model." IEEE Transactions on Vehicular Technology. Vol. 53, pp. 1586-1593, 2004.

[10] Y. Gao, X. Zhang, J. Yang, and B. Guo, Estimation of State-of-Charge and State-of-Health for Lithium-Ion Degraded Battery Considering Side Reactions. Journal of The Electrochemical Society, 2018, vol. 165, pp. A4018-A4026.

[11] L. Pei ,C. Zhu, Chunbo ,T. Wang, R Lu , Chan, C.C., "Online peak power prediction based on a parameter and state estimator for lithium-ion batteries in electric vehicles". Energy, Elsevier, vol. 66(C), p 766-778, 2014.

[12] S. Xiang, G. Hu, R. Huang, F. Guo, and P. Zhou, Lithium-Ion Battery Online Rapid State-of-Power Estimation under Multiple Constraints. Energies 2018, vol. 11, pp. 283.

[13] B. Wang, X. Peng and D. Liu, "Airborne Sensor Data-Based Unsupervised Recursive Identification for UAV Flight Phases," IEEE Sensors Journal, vol. 20, no. 18, pp. 10733-10743, 15 Sept.15, 2020,

[14] D'informatique. D, N. Ese, Esent. Pr, Au. Ee, Gers. Felix, Hersch. Prof, Esident. Pr, and Frasconi. Prof, Long Short-Term Memory in Recurrent Neural Networks, May 2001, pp. 2366.

[15] D. Liu, L. Li, Y. Song , L. Wu, and Y. Peng, Hybrid state of charge estimation for lithium-ion battery under dynamic operating conditions. International Journal of Electrical Power & Energy Systems, 2019, vol. 110, pp. 48-61.

[16] J. Zhang, P. Wang, R. Yan, and R. Gao, Long short-term memory for machine remaining life prediction. Journal of Manufacturing Systems. 2018, vol. 05, pp. 011.

[17] F. Yang, S. Zhang, W. Li, and Q. Miao, State-of-charge estimation of lithium-ion batteries using LSTM and UKF. Energy, 2020, pp. 117664.

[18] F. Sun, R. Xiong, H. He, and W. Li, Aussems, J.E.E. Model-based dynamic multi-parameter method for peak power estimation of lithium-ion batteries. Applied Energy. 2012 , vol. 96, pp. 378-386.

[19] Graves, Alex. Generating Sequences With Recurrent Neural Networks. ArXiv abs/1308.0850, 2013.

UAV Anomaly Detection Using Active Learning and Improved S3VM Model

Dawei Pan*
College of Information and Communication Engineering
Harbin Engineering University
Harbin, China
pandawei@hrbeu.edu.cn

Longqiang Nie
College of Information and Communication Engineering
Harbin Engineering University
Harbin, China
nielongqiang@hrbeu.edu.cn

Weixin Kang
College of Information and Communication Engineering
Harbin Engineering University
Harbin, China
kangweixin@hrbeu.edu.cn

Zhe Song
College of Information and Communication Engineering
Harbin Engineering University
Harbin, China
songzhe1926@hrbeu.edu.cn

Abstract—**Unmanned aerial vehicles can complete various specific tasks and play an increasingly important role in various fields. However, UAVs often have anomalies during flight, which is likely to cause huge losses. Therefore, the use of data-driven anomaly detection methods has attracted attention. Because UAV sensors have less labeled data and more unlabeled data, semi-supervised support vector machine (S3VM) classification method is introduced. Considering that unlabeled instances may not have the information content and the distribution assumptions used may not reflect the real time series distribution well, the pure S3VM classification method may not necessarily achieve the ideal classification effect. For existing time series data, active sampling is used to mine the most valuable data of the classifier model through margin sampling (MS). In addition, due to the overlapping of different types of data, the classification boundary of the S3VM classifier cannot be located in a low-density region. By adjusting the distance of most samples to the classification boundary, the classification boundary is in a low-density region, which meets the theoretical design requirements of the classifier. Experiments show that combining MS active learning and improved S3VM classification method will achieve the ideal classification effect. Using the UAV time series data, according to the classification prediction results of the classifier model during the operation stage of the training data training, an estimate of the prediction uncertainty is given. Anomaly detection is realized by comparing the predicted value with the uncertainty interval for classification. This paper uses three sets of UAV channel telemetry data for experimental verification. Using MS active learning and the improved S3VM algorithm compared with SVM and S3VM algorithms, it verifies the effectiveness of the algorithm in different UAV data sets for anomaly detection.**

Keywords—UAV, active learning, S3VM, time series, anomaly detection

I. INTRODUCTION

As the application of UAV is becoming more and more extensive, the corresponding functions and technical indicators are also continuously improved, and the flight safety of UAV is also paid more and more attention. Once the sensor of the UAV flight control system fails, the wrong information will be fed back to the flight control computer, which will cause serious consequences such as abnormal flight of the UAV [1]. Therefore, the use of abnormal data from UAV sensors for fault detection and diagnosis is the research direction of this subject and has very important practical significance.

In the field of anomaly detection, a series of the most advanced anomaly detection methods have been proposed, mainly divided into three categories: knowledge-based methods, model-based methods and data-driven methods. Knowledge-based methods can obtain anomaly patterns by summing up the experience of experts in specific fields, and establish corresponding anomaly detection systems, and eventually achieve the same anomaly detection pattern [2]. Model-based methods usually construct an observer system model by establishing an accurate physical model, and then compare the observer or filter whose estimated value has the actual measured value. According to the residuals generated, anomalies are detected [3]. The data-driven method automatically learns the process data of system behavior simulation based on the set system [4]. The sensor data generated during the operation of the UAV is a typical time series. Combining the characteristics of the actual UAV historical data with less labeled data and a large amount of unlabeled data, the data-driven method is mainly applied to the UAV transmission abnormal situation of sense data.

Anomaly detection methods based on prediction are one of the data-driven anomaly detection methods. In recent years, they have become a research hotspot, especially in the application of time series anomaly detection. The performance of these methods depends on accurate prediction models. Common prediction methods include supervised learning classification represented by SVM, unsupervised learning clustering represented by K-means and LSTM Neural Networks [5-7] and so on.

Active learning (AL) as a new machine learning method, its main goal is to effectively find samples with a large amount of information in the training data set, reduce the number of training data, and efficiently train the classification model [8]. In the process of model training, samples that have no obvious effect on the classification performance of the classifier are ignored, thereby greatly reducing the cost of labeling and improving the quality of the labeled sample set. The performance of active learning depends on the quality of the sampling strategy. Different sampling strategies have different effects on improving the performance of the current classifier.

Semi-Supervised Learning (SSL) uses a small number of labeled samples and a large number of unlabeled samples to update the training classifier. Compared with active learning methods, semi-supervised learning further reduces the cost of manual labeling and saves time and effort. However, the simple semi-supervised classification model is greatly affected by the initial training data, and incorrect predictions of labeled and unlabeled samples will lead to classification model classification accuracy. In this paper, semi-supervised learning uses semi-supervised support vector machine (S3VM). Considering that samples of different categories in time series data may overlap, the classification boundary of the classification model is actually not in the low-density area required by the theory. Adjust the distance between the labeled sample and the classification boundary so that the adjusted classification boundary is in a low-density area. The adjusted S3VM classification model has better classification performance than before adjustment.

The algorithm of the thesis is divided into two parts. In the active learning part, the samples with rich information and significant differences are searched from the low-density areas of the samples for labeling to improve the quality of the labeled data. In the semi-supervised learning part, the active learning idea is used to query samples that are between the current classifier hyperplane and contain a certain amount of information and credibility from unlabeled samples, and use the adjusted S3VM to train the classifier. Active learning and semi-supervised learning have good complementarity. Combining them effectively can not only improve the performance of classification models, but also reduce the workload in the formation of classification models. In order to obtain higher accuracy of anomaly classification of UAV sensor data, this paper proposes an anomaly detection algorithm based on active learning and improved S3VM, including data label processing, anomaly detection network model establishment and anomaly detection evaluation indicators. Simulation data was verified using UAV sensor data and compared with SVM and S3VM algorithms.

II. INTRODUCTION TO BASIC ALGORITHMS

A. Active Learning Algorithm

For supervised learning models, enough labeled samples are a prerequisite for obtaining a high-precision classifier. With the rapid development of sensor technology, data collection is getting easier. At the same time, there is less labeled data and more unlabeled data, and manual labeling of data is expensive.

As shown in the Figure 1, in pool-based active learning, there is a training set L composed of labeled samples, and an unlabeled data pool U. The classifier uses the samples in the training set L for training. Through a specific point selection strategy Q, a part of the valuable sample points is selected from U to form a subset U_0. The expert completes the labeling of U_0 and adds the sample points in U_0 together. The labels are added to the labeled sample set L together for the next iteration, and the point is removed from U.

This paper proposes to use active learning and selective learning based on data to improve the income of manual labeling. The improvement of entity labeling model based on semi-supervised learning method aims to use unlabeled data to further improve the performance of the model under the premise of making full use of labeled data.
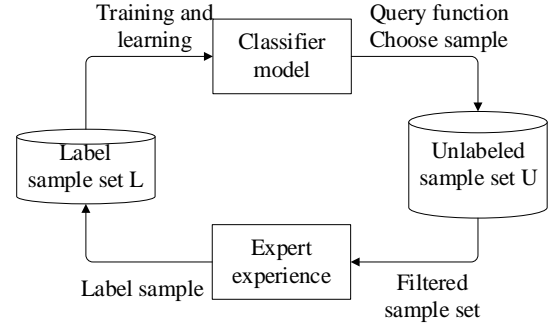
Fig. 1. The process of implementing active learning.

Margin sampling (MS) is an active learning algorithm mainly used to support vector machine classification models. The distance can intuitively reflect the classification confidence of the point, and the distance between the unlabeled data point and the classification hyperplane is calculated, the closer it is, the lower the confidence of the classification model on the sample [9]. In edge sampling, the algorithm's point selection strategy is to select the sample point closest to the classification interface, because for the classifier, such a sample has a large amount of classification information, and the more likely it is to become the next support vector.

Given an already labeled training set:

$$L = \left\{ (x_1, y_1), ..., (x_m, y_m) \right\}, x_i \in R^d \tag{1}$$

Corresponding class tag $y_i \in \{-1, 1\}$. Considering the classification problem, the distance from any sample point to the SVM classification surface can be expressed as:

$$f(x_i) = \sum_{j=1}^{n} p_j y_j G(x_j, x_i) + q \tag{2}$$

In this formula, p_j is a non-zero support vector coefficient, y_j is the classification mark $y_i \in \{-1, 1\}$, $G(x_j, x_i)$ is the kernel function, which is used to calculate the similarity between the candidate example x_i and the support vector x_j, q is Real rational number.

The goal of MS is to find the sample point closest to the interface from a large number of unlabeled sample sets U (there are n unlabeled samples, and n>>m). The conditions satisfied by this sample point can be as follows:

$$x^{MS} = \arg \min_{x_i \in U} \{ \min_z | f(x_i, z) | \} \tag{3}$$

B. S3VM Algorithm

Among the many semi-supervised learning methods, S3VM is very popular. It is based on the clustering hypothesis and attempts to standardize and adjust decision boundaries by exploring unlabeled data. In order to utilize unlabeled data, we need to add two restrictions on unlabeled data points based on the original SVM. One limitation is to assume that the unmarked point belongs to the category, and then calculate its misclassification rate; another limitation is to assume that this point belongs to the category, and also calculate its misclassification rate.

The SVM in supervised learning attempts to find a partitioned hyperplane that maximizes the interval between the support vectors on both sides, which is the idea of "maximum partition interval" [10]. For semi-supervised learning, S3VM considers that the hyperplane needs to pass through areas with low data density. The core idea is to try to find a suitable label assignment for unlabeled samples, so as to maximize the interval after the hyperplane is divided. S3VM adopts a local search strategy for iterative solution, that is, an initial SVM is first trained using the labeled sample set, and then the classification model is used to mark unlabeled samples, so that all samples are labeled and based on these labeled samples are retrained SVM, and then look for error-prone samples and continue to adjust.

For non-linear data sample classification, the SVM hyperplane must meet all sample classification requirements, so the following formula is met:

$$y_i\left[\left(w \cdot x_i + b\right)\right] - 1 + \xi_i \geq 0, i = 1, 2, \ldots, n \quad (4)$$

The best classification boundary in the SVM algorithm is to take the minimum value of the following formula.

Where C is a constant that controls the degree of penalty of the sample, which needs to be greater than 0.

$$\frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{n} \xi_i\right) \quad (5)$$

In the non-linear sample division, we usually turn it into a linear problem in high dimensions for analysis, in order to solve the case of non-classifiable samples, and in order to avoid the "dimension of dimensionality", we introduced kernel functions. The kernel function is used as the inner product function. Through functional theory, we can know that as long as the function meets the Mercer condition, then we can use this function as the inner product function, so in the case of processing that cannot be divided, we first map the data to the kernel function to high-dimensional space, thus solving its classification problem.

According to the characteristics of the algorithm, the Gaussian kernel is selected here:

$$k\left(x_1, x_2\right) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2\delta^2}\right) \quad (6)$$

The objective function of S3VM:

$$\min_{w,b,\hat{y},\omega} \frac{1}{2}\|w\|_2^2 + C_l \sum_{i=1}^{l} \xi_i + C_u \sum_{i=l+1}^{m} \xi_i \quad (7)$$

$$y_i(w^T x_i + b) \geq 1 - \xi_i, i = 1, 2, \ldots, l,$$

$$\hat{y}_i(w^T x_i + b) \geq 1 - \xi_i, i = l+1, l+2, \ldots, m,$$

$$\xi_i \geq 0, i = 1, 2, \ldots, m,$$

S3VM looks for a decision boundary in the low-density interval of unlabeled data. In fact, it can be seen that the objective equation of S3VM is to move the unlabeled data as far as possible from the decision boundary (the decision boundary passes through the low level of unlabeled data as much as possible density area). The assumption of S3VM is that the sample classes can be well separated, and the decision boundary falls in the low-density area of the sample feature

space and does not pass through dense unlabeled data. If this assumption is not satisfied, S3VM may cause poor performance.
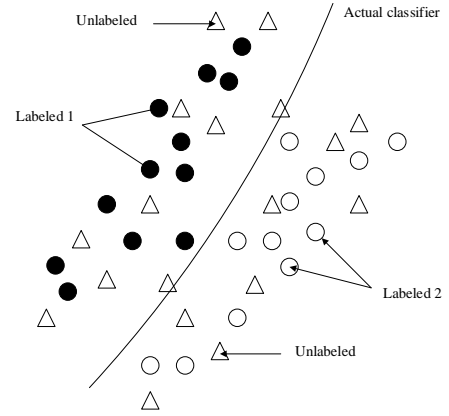
Fig. 2. The semi-supervised classification boundary is in a low-density area.

C. Improved S3VM Algorithm

Semi-supervised classification methods may even achieve unsatisfactory results. Consider the possible reasons. The first is that unlabeled instances may be unreliable, which may mislead the classification. The second is that the distribution assumptions used may not reflect the real data well distribution, which may mislead the classification.

For unlabeled instances that may be unreliable, an active learning algorithm is used to find the mostly worthy samples in a large number of unlabeled data pools by using strategies. For the distribution hypothesis that may not reflect the true data distribution well, we consider further focusing on the cluster hypothesis, which specifies that the decision boundary falls in the low-density area of the sample and does not pass through dense unlabeled data, that is, the classification passes through the low-density area.

When the samples of different categories are severely overlapped, the actual boundary is not in the low density area. At this time, first calculate the distance between each labeled instance and the boundary, and then include the cluster boundary in the instance of S3VM. In this way, the classification boundary is guided through the cluster boundary, which is in the low-density area. At the same time, when the classes or clusters do not overlap significantly, the true boundary is in a low-density area. In addition, the determination of the boundary actually describes the confidence of the classification, so the new method takes into account the confidence of unlabeled instances in the classification.

Introduce a distance vector describing the distance between individual instances and the classification boundary, $V \in Q_n$. Each element of the distance vector is expressed as:

$$V_i = \left|d(x_i - v_1) - d(x_i - v_2)\right| \quad (8)$$

When V_i is larger, the instance is farther away from the cluster boundary, otherwise, it is more likely to be within the cluster boundary. Assuming that the instances within the cluster boundary are also within the class boundary.

For the assumption-based method, it is assumed that the decision boundary is not in the low-density area [11]. By adding a cluster structure to the improved S3VM, the decision

boundary can be adjusted to the real boundary. At the same time, when the classes or clusters do not overlap significantly, the true boundary is indeed in a low-density area. In this case, the detected classification boundary will be a low-density separation. If there is no adjustment, because the boundary is in the low density area. This undoubtedly alleviates the effect of inaccurate distribution assumptions to a certain extent. Therefore, it can also mitigate the effect of inaccurate distribution assumptions and improve the reliability of semi-supervised classification.

For a fixed $\hat{y}_j$, the optimization problem of minimum value $f(x)$ is expressed as follows, Using an iterative strategy to solve:

$$f(x) = \sum_{i=1}^{n} \partial_i^H (x_i, x) \tag{9}$$

Lagrange operator:

$$\partial = \left[a_1, a_2 ... a_n \right] \tag{10}$$

H is matrix: $H = \left[H_l H_u \right]$ (11)

$$H_{ll} = \left\langle \theta(X_l), \theta(X_l) \right\rangle_h \tag{12}$$

$$H_{lu} = \left\langle \theta(X_l), \theta(X_u) \right\rangle_h \tag{13}$$

At the same time, with fixed $f(x)$, the optimization problem $\hat{y}_j$ can be written as:

$$\min_{\hat{y}_j} \sum_{j=l+1}^{n} V_j \left(f(x_j) - \hat{y} V_j \right)^2 \tag{14}$$

Because $V_j \geq 0$, so for each x_j, if $f(x_j) \geq 0$, $\hat{y}_j = 1$, otherwise $\hat{y}_j = -1$.

III. ANOMALY DETECTION MODEL DESCRIPTION

A. Model Description
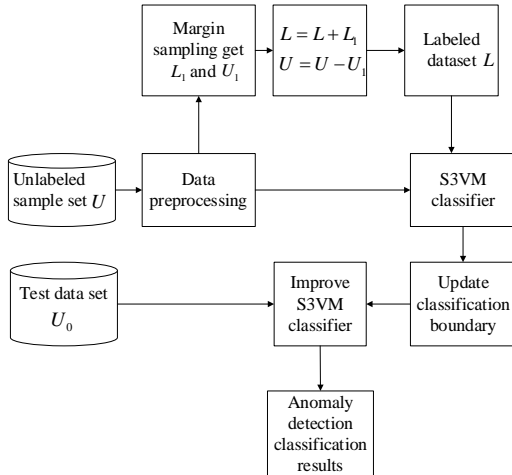
Fig. 3. Anomaly detection framework.

The model is designed from two aspects. First, the selected samples for manual labeling must be the most informative for the current classification model to maximize the efficiency of manual labeling; second, for the information contained in the remaining large number of unlabeled samples, the semi-supervised classification algorithm should be further utilized. So we proposed the algorithm in this paper, which mainly includes several key steps such as initial sample selection and classifier model training, MS active learning, S3VM semi-supervised learning, and sample boundary update. The complete algorithm framework is shown in the Figure 3.

The anomaly detection model based on MS active learning and improved S3VM is described as follows: initially, the total sample set is mainly composed of two parts, a small number of labeled samples and a large number of unlabeled samples. First, the support vector machine is trained through the labeled sample set, and an initial classification model can be obtained. Using the existing classification model to predict the category of the unlabeled sample, the S3VM classification model is obtained. The goal of MS active learning is to find the sample point with the closest distance to the interface from a large number of unlabeled samples, and based on the selection index for comprehensive evaluation of sample representativeness and information degree, select multiple most use valuable samples for labeling, and iteratively add the labeled samples to the labeled sample set to form a new and expanding training sample set. A new S3VM classification model is trained using the expanded labeled sample set.

For the obtained S3VM classification model, calculate the distance between each labeled instance and the boundary, and then include the cluster boundary in the instance of S3VM. In this way, the classification boundary is guided through the cluster boundary, so that the classification boundary is in a low-density area, updated more reliable S3VM classification model.

B. Data Preprocessing

There are isolated sample points and noises in the original UAV telemetry data. Therefore, before performing UAV flight status recognition, the data needs to be properly preprocessed to remove isolated sample points. In this paper, the principle 3σ is used to process the first-order difference of the data, and the isolated sample points are eliminated. Suppose that in a certain flight, parameter x records the value $\{x_1, x_2, ..., x_n\}$ of n moments in total, first calculate the first-order difference y, $y_i = x_{i+1} - x_i$, and $i = 1, 2, ..., n-1$ of the parameter. Calculate the average $\bar{y}$ and standard deviation σ of the first-order difference y, and delete the original data corresponding to the data points other than 3σ according to the $[\bar{y} - 3\sigma, \bar{y} + 3\sigma]$ principle.

UAV telemetry data contains noise introduced in the measurement process, this paper uses a moving average method to achieve smooth filtering of the data. The core idea is to calculate the moving average of multiple sequential values of the sequence, thereby forming a new sequence of average values, that is, the value of each point in the sequence is replaced by the subsequent sequence average [12]. High-frequency noise is filtered out, and the data changes more smoothly. The process is as follows:

$$\Delta y_s = \left\{ \Delta y_s{}^1, \Delta y_s{}^2, ..., \Delta y_s{}^t \right\} \quad (15)$$

$$\Delta y_s{}^t = \partial \left(\Delta y_s{}^{t-1} - \Delta y_s{}^t \right) + \Delta y_s{}^t \quad (16)$$

C. Model Steps

The specific steps of the anomaly detection algorithm are as follows:

- Input: labeled sample set L, unlabeled sample set U, the number of samples actively selected during each round of iteration, the number of labeled samples, the number of labels for each sample in S3VM, and the S3VM related parameters.

- Output: Classification results for the current improved S3VM classification model.

- Iteration process:

1) In MS active learning, use formula (1) to calculate the distance from the sample point to the classification hyperplane. The closer the sample is to the classification interface, the lower the confidence of the classification model. For the classification interface, the sample has a large amount of classification information. Find the most representative and informative sample and mark it so that the marked sample is formed among the unmarked samples.

2) Active learning is an iterative method. In this iterative process, the training process of active learning only needs to make the classifier reach the expected classification accuracy rate. It is not necessary to supplement the samples to continue training, and iterative stops.

3) Using formulas (2) - (5), use the labeled samples generated by MS active learning and the original unlabeled samples to train the S3VM to obtain the initial classification model.

4) Using the formula of the distance between the marked instance and the boundary, adjust the clustering center to update the classification boundary through formulas (6) - (8), and then include the cluster boundary of the sample category in the S3VM. The classification boundary is guided through the cluster boundary, making the classification boundary in a low-density area, the updated S3VM classification model is more reliable.

5) Use the test data set (unlabeled samples) as the input of the improved S3VM classifier that has been trained to perform classification and recognition, thus completing the abnormal classification of different sensor time series.

IV. EXPERIMENTAL RESULTS

A. Experimental Data

In order to fully verify the performance proposed based on active learning and improved S3VM. The verification data in this paper uses multiple sets of telemetry data from UAV. Compared with the abnormal detection methods based on SVM and S3VM, the performance of the abnormal detection method based on active learning and S3VM method in anomaly detection of UAV data points is divided into pre-processed data set into training set and test set 60% is used as the training set, and the remaining 40% is used as the test set .

In this paper, three sets of telemetry data are randomly selected as experimentally verified data sets. In the simulation

experiment, the accuracy of classification prediction and the detection of abnormal points are performed after the training is completed [13-14].

TABLE I. EXPERIMENTAL DATA SETTING

Sample name	*Shinto 1*	*Shinto 2*	*Shinto 3*
Total number of samples	42	45	45
Number of features	10	10	10
Point anomaly sequence number	15	18	10

Because the total amount of telemetry data of each group of sequences ranges from 300 to 8 000 points, the telemetry data is divided into two parts, of which 60% is the training set and the rest 40% is used for testing. For the three sets of selected data sets, each set of data for labeled data and unlabeled data, for each set of labeled data, the number of labeled samples is quantitatively divided into 5, 10, combined with the same number of unlabeled data, the prediction accuracy of unlabeled instances is quantitatively obtained through experiments, and the anomaly detection algorithm is obtained in The effect of abnormal sensor classification.
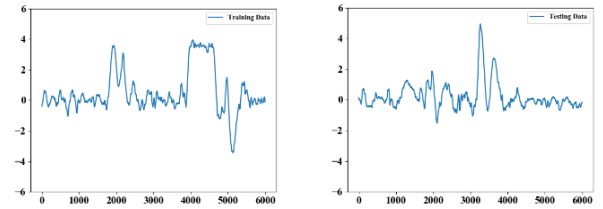
Fig. 4. Training data and test data

B. Experimental Performance Index

The abnormal test results are generally divided into four different types: true positive, false positive, false negative and true negative, as shown in Table 2. In order to evaluate the fault detection performance of the fault detection method, two methods of true positive rate (TPR) and false positive rate (FPR) were used in the experiment. The larger TPR value and the smaller FPR value represent the satisfactory detection performance of this detection method [15].

TABLE II. TYPES OF ANOALY DETECTION

Detection Results	Normal Data	Faulty Data
Normal	True Positive (TP)	False Positive (FP)
Faulty	False Negative (FN)	True Negative (TN)

$$TPR = \frac{(Num)TP}{(Num)TP + (Num)FN} \quad (17)$$

$$FPR = \frac{(Num)FP}{(Num)FP + (Num)FN} \quad (18)$$

C. Analysis of Results

By comparing with the SVM and S3VM-based anomaly detection methods, the performance of anomaly detection

based on active learning and improved S3VM anomaly detection methods in UAV data points can be seen, whether it is the TPR indicator or the FPR indicator, they have reached the most results. The detection method has better detection performance.

TABLE III.   COMPARISON OF INDICATORS OF DIFFERENT MODELS

Number of labeled samples 5		
Model	TPR	FPR
SVM	76.3%	13.1%
S3VM	80.5%	12.5%
Article model	90.8%	9.8%
Number of labeled samples 10		
Model	TPR	FPR
SVM	89.3%	9.2%
S3VM	81.2%	15.9%
Article model	92.7%	8.2%

Each data set is randomly divided into two parts, one for training and the other for training and the training set contains 5 and 10 labeled samples respectively. This process is repeated 20 times as the classifier learns, and the average accuracy and variance are reported. The Gaussian kernel width parameter used in the algorithm model is the average value of the distance between the selected instance pairs, and the average prediction performance of the unlabeled data can be obtained.

V. CONCLUSION

This paper proposes an anomaly detection model based on active learning and improved S3VM classification. The model comprehensively considers the information and representativeness of unlabeled samples, and eliminates outlier samples. By adding MS active learning, the classification accuracy of the classifier can be improved by marking a small number of samples. In S3VM, by adjusting the distance between the sample instance and the classification boundary, the original sample cluster boundary is changed. The updated cluster boundary expresses the classification boundary, so that the adjusted S3VM classification boundary is in the low-density area of unlabeled data. The obtained S3VM algorithm will have relatively excellent classification performance. Use the drone sensor data set for training and testing to obtain a new anomaly detection model, which is compared with SVM and S3VM. The performance indicators of the experiment are reflected by the true positive rate (TPR) and false positive rate (FPR). The TPR value and the smaller FPR value reflect that the model has better anomaly classification and detection performance. Through comparison, it is found that not only the effectiveness of the model is verified, but also a larger TPR value and a smaller FPR value are obtained. The most satisfactory anomaly classification detection performance is obtained among the three algorithms. In the future, a variety of abnormal modes will be used to verify the proposed method and try to conduct more comprehensive comparative experiments with other methods.

REFERENCES

[1] ALJANAIDEH K F, BERNESTEIN D S. Aircraft sensor health monit ring based on transmissibility operators[J]. Journal of Guidance, Control, and Dynamics, 2015, 38(8):1492-1495.

[2] E. F. Carrasco, J. RodríGuez, A. Puñal, E. Roca, and J. M. Lema, "Rule-based diagnosis and supervision of a pilot-scale wastewater treatment plant using fuzzy logic techniques," Expert Systems with Applications,vol. 22, pp. 11-20, 2002.

[3] NA R ASIMHAN S , B R OWNSTON L. HyDE‐A general framework for stochastic and hybrid model-based diagnosis[J]. Proceedings 18th Interational. Workshop Principles Diagnosis DX-07, 2007: 162-169.

[4] Z. Birnbaum, A. Dolgikh, V. Skormin, E. O'Brien, D. Muller, and C.Stracquodaine, "Unmanned Aerial Vehicle Security Using Recursive Parameter Estimation," Journal of Intelligent & Robotic Systems, vol.84, pp. 107-120, 2016.

[5] SMA R T E, B R OWN D, DENMAN J. Combining multiple classifiers to quantitatively rank the impact of abnormalities in flight data[J]. Applied Soft Computing, 2012, 12(8): 2583-2592.

[6] LI L, DAS S, JOHN HANSMAN R, et al. Analysis of flight data using clustering techniques for detecting abnormal operations ［J］. Journal of Aerospace information systems，2015，12(9): 587-598.

[7] Ergen Tolga,Kozat Suleyman Serdar. Unsupervised Anomaly Detection With LSTM Neural Networks.[J]. IEEE transactions on neural networks and learning systems,2019.

[8] TONG S, KOLLER D. Support vector machine active learning with applications to text classification [J]. Journal of Machine Learning Resrarch,2012,2(1):45-66.

[9] LIU K, QIAN X, WANG Z Q. Overview of active learning algorithms[J]. Computer Engineering and Applications, 2012, 48(34):1-4. DOI:10.3778/ j.issn.1002-8331.1205-0149.

[10] ZHOU Zhi-hua. Machine Learning[M]. Beijing: Tsinghua University Press, 2016.

[11] Li Y-F, Zhou Z-H. Towards making unlabeled data never hurt. IEEE Trans Pattern Anal Mach Intell. 2015, 37(1):175–188.

[12] Dong J Y, Pang J Y, et al. Anomaly detection method of spacecraft telemetry data integrated with LSTM[J]. Chinese Journal of Scientific Instrument, 2019, 40(07): 22-29.

[13] Aboutalebi, P.; Abbaspour, A.; Forouzannezhad, P.; Sargolzaei, A. A Novel Sensor Fault Detection in an Unmanned Quadrotor Based on Adaptive Neural Observer. J. Intell. Robot. Syst. 2018, 90, 473–484.

[14] Sun, R.; Cheng, Q.; Wang, G.; Ochieng, W.Y. A Novel Online Data-Driven Algorithm for Detecting UAV Navigation Sensor Faults. Sensors 2017, 17, 2243.

[15] B. Wang, Z. Wang, L. Liu, D. Liu and X. Peng, "Data-Driven Anomaly Detection for UAV Sensor Data Based on Deep Learning Prediction Model," 2019 Prognostics and System Health Management Conference (PHM-Paris), Paris, France, 2019, pp. 286-290.

Research on Multivariate Variational Mode Decomposition Method and Its Application to Bearing Fault Diagnosis

Qiuyu Song
School of Rail Transportation
Soochow University
Suzhou, China
749875042@qq.com

Xingxing Jiang*
School of Rail Transportation
Soochow University
Suzhou, China
jiangxx@suda.edu.cn

Jun Wang
School of Rail Transportation
Soochow University
Suzhou, China
junking@suda.edu.cn

Changqing Shen
School of Rail Transportation
Soochow University
Suzhou, China
cqshen@suda.edu.cn

Juanjuan Shi
School of Rail Transportation
Soochow University
Suzhou, China
jshi091@suda.edu.cn

Weiguo Huang
School of Rail Transportation
Soochow University
Suzhou, China
wghuang@suda.edu.cn

Zhongkui Zhu
School of Rail Transportation
Soochow University
Suzhou, China
zhuzhongkui@suda.edu.cn

ABSTRACT—The bearing failure diagnosis methods based upon variational mode decomposition (VMD) have been researched extensively in recent years. However, these methods are only capable of dealing with single channel data, in which the amount of information is limited and the anti-interference ability needs to be further enhanced. Recently, as one of extensions of VMD, the multivariate variational mode decomposition (MVMD) was put forward to deal with the multivariate signal. For the sake of monitoring the operation state of equipment more comprehensively and diagnose mechanical failure more accurately, we propose a new bearing failure diagnosis method based upon MVMD. In particular, to select a suitable value of decomposed modes K, a scheme which combines correlation coefficient and kurtosis criterion is innovatively proposed to enhance the performance of the MVMD. The result of experimental verification has demonstrated the enhanced performance of this new method in transient extraction and faults detection of bearings over the original VMD method.

Keywords—multivariate variational mode decomposition, fault diagnosis, correlation coefficient, kurtosis

I. INTRODUCTION

Nowadays, rotating machinery has been widely used in various applications, especially in the field of aeronautic industry. It usually works under non-ideal conditions, such as the high-speed, severe pollution and heavy load. Furthermore, as a key component of aero-engine, rolling bearing is of particular importance to the operation efficiency and reliability of the aero-engine. Hence, realizing the bearing failure identification of local damage accurately and reliably is the key to enhance the mechanical performance and ensure the safety operation [1-3].

In recent decades, a great variety of techniques have been put forward to process the fault signal, such as time-frequency analysis [4-6], envelope demodulation [7], deconvolution-based methods [8], adaptive mode decomposition [9] and so on. Compared with other methods, the adaptive mode decomposition shows advantages in focusing on mapping arbitrary complex multi-component signals to single-component signals and revealing weak transient characteristics without prior transmission. Based on this, a wide range of adaptive mode decomposition techniques, including local mean decomposition (LMD) [10], empirical mode decomposition (EMD) [11,12], empirical wavelet transform (EWT) [13,14], variational mode decomposition (VMD) [15,16], and their variants have become the hot topic for bearing fault diagnosis.

However, most of the adaptive mode decomposition methods are mainly suitable for handling single channel data which involves rather limited information. Considering that the information contained in the multi-channel data is more complete and in line with actual engineering applications, multi-channel signal processing technology has become the focus of current research. Rehman, et al. first proposed the fully multivariate extension of VMD, i.e., multivariate variational mode decomposition (MVMD) [17]. Although MVMD succeeds all the merit of the original VMD, it is still trapped by the problem of initial parameter selection, such as the value of decomposed modes K. If the value of K is chosen improperly, the final decomposition will be affected, and even result in inaccurate analysis.

In this paper, to address this problem and utilize MVMD in bearing fault diagnosis, a reliable and efficient bearing failure diagnosis method based on MVMD is presented completely and clearly. Particularly, the major devotions and advantages of the new method are summarized into three aspects. (1) This new method first applies the MVMD algorithm to the field of mechanical failure diagnosis. Based on the basic principle of MVMD, a fault characteristic extraction method for rolling bearing is successfully established. (2) By combining correlation coefficient and kurtosis criterion, the proposed method provides a scheme for selection of parameter K. When the number of decompositions is difficult to determine, the kurtosis of the maximum correlation component under different values of K are compared. Then, the K value that produces the largest kurtosis should be determined as the optimal solution. (3) The analysis of experiment data for faulty bearing acquired by

---

* Xingxing Jiang is the corresponding author. (e-mail: jiangxx@suda.edu.cn).

an aero-engine simulator confirmed the unique advantages of this new method in the application of bearing failure diagnosis. This new method has a good ability to extract the fault components in each channel accurately and efficiently, thus suitable for fault diagnosis.

II. MECHANICAL FAULT DIAGNOSIS METHOD BASED ON MVMD

A. Multivarite Variational Mode Decomposition

1) Establishment of the MVMD Model

MVMD is appreciated for its novel expansion of the traditional VMD algorithm for multichannel data sets, and is utilized for the purpose of extracting K meaningful modes $\mathbf{u}_k(t)$ from an input real-valued signal $\mathbf{x}(t)$ containing C channels, i.e, $\mathbf{x}(t) = \left[x_1(t), x_2(t), \cdots, x_C(t) \right]$

$$\mathbf{x}(t) = \sum_k^K \mathbf{u}_k(t) \tag{1}$$

where, $\mathbf{u}_k(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \dots \\ u_C(t) \end{bmatrix} = \begin{bmatrix} a_1(t)\cos(\phi_1(t)) \\ a_2(t)\cos(\phi_2(t)) \\ \dots \\ a_C(t)\cos(\phi_C(t)) \end{bmatrix}$, $a_i(t)$ and

$\phi_i(t)$ represent the amplitude and phase function corresponding to the ith signal component respectively.

This MVMD algorithm adaptively decomposes multi-channel signals into several multivariate modulated oscillations $u_{k,c}(t)$ with definite physical meaning under the condition that each component is tightly supported on its own central frequency ω_k. The sum of the bandwidth of all modal components ought to be minimized. In the iterative process, the center frequency and the component bandwidth are updated for the real mode component. The steps of the algorithm are described as below.

a) After the Hilbert transform, the analytic signal of $u_k(t)$ is obtained:

$$\mathbf{u}_+^k(t) = \mathbf{u}_k(t) + j\mathcal{H}\mathbf{u}_k(t) = \begin{bmatrix} u_+^1(t) \\ u_+^2(t) \\ \dots \\ u_+^c(t) \end{bmatrix} = \begin{bmatrix} a_1(t)e^{j\phi_1(t)} \\ a_2(t)e^{j\phi_2(t)} \\ \dots \\ a_c(t)e^{j\phi_c(t)} \end{bmatrix} \tag{2}$$

b) According to the character of frequency shift, the spectrum of each analytic signal is converted to baseband by setting the initial center frequency ω_k:

$$e^{j\omega_k t}\mathbf{u}_+^k(t)$$

c) The gradient form of the squared L2-norm of this shifted signal $\mathbf{u}_+^k(t)$ is used to assess the bandwidth of $\mathbf{u}_k(t)$, and then the obtained cost function for MVMD is given:

$$f = \sum_k \sum_c \left\| \partial_t \left[u_+^{k,c}(t) e^{-j\omega_k t} \right] \right\|_2^2 \tag{3}$$

In Eq. (3), $u_+^{k,c}(t)$ represents the analytical modulation signal corresponding to mode k and channel c.

d) now the constrained optimization problem for MVMD is presented:

$$\min_{\{u_{k,c}\},\{\omega_k\}} imize \left\{ \sum_k \sum_c \left\| \partial_t \left[u_+^{k,c}(t) e^{-j\omega_k t} \right] \right\|_2^2 \right\}, \tag{4}$$
$$s.t. \sum_k u_{k,c}(t) = x_c(t), c = 1,2\cdots C$$

2) Solution of the MVMD Model

There are two penalty terms used to solve this variational constraint object function. On the one hand, a quadratic term is used to enhance the reconstruction fidelity, and on the other the term with Lagrangian multipliers λ is introduced to make sure of the strict satisfaction of constraints. Then an resulting unconstrained variational problem is obtained and ensures the solution procedure faster. As a result, final augmented Lagrangian function is obtained as below.

$$L\left(\{u_{k,c}\},\{\omega_k\},\lambda_c\right) = \alpha \sum_k \sum_c \left\| \partial_t \left[u_+^{k,c}(t) e^{-j\omega_k t} \right] \right\|_2^2 \tag{5}$$
$$+ \sum_c \left\| x_c(t) - \sum_k u_{k,c}(t) \right\|_2^2 + \sum_c \left\langle \lambda_c(t), x_c(t) - \sum_k u_{k,c}(t) \right\rangle$$

Then, this problem is solved cleverly with the help of the alternate direction method of multipliers algorithm (ADMM)[18]. This ADMM algorithm transforms a fairly complex optimization problem into several relatively simpler sub-optimization problems.

- Update mode

The problem related to the mode update is given priority to emphasize, such as the following sub-optimization problems in the nth iteration:

$$u_{k,c}^{n+1} \leftarrow \arg\min_{u_{k,c}} \mathcal{L}\left(\left\{u_{i<k}^{n+1}\right\}, \left\{u_{i\geq k}^n\right\}, \left\{\omega_i^n\right\}, \lambda^n\right) \tag{6}$$

Eq. (6) is also equivalent to the following formula:

$$u_{k,c}^{n+1} = \arg\min_{u_{k,c}} \left\{ \alpha \left\| \partial_t \left[u_+^{k,c}(t) e^{-j\omega_k t} \right] \right\|_2^2 + \left\| x_c(t) - \sum_i u_{i,c}(t) + \frac{\lambda_c(t)}{2} \right\|_2^2 \right\} \tag{7}$$

According to Parseval Theorem, the algorithm is converted to frequency domain:

$$\hat{u}_{k,c}^{n+1} = \arg\min_{u_{k,c}} \left\{ \alpha \left\| j\omega\left[1 + \text{sgn}(\omega+\omega_k)\hat{u}_k(\omega+\omega_k)\right] \right\|_2^2 + \left\| \hat{x}_c(\omega) - \sum_i u_{i,c}(\omega) + \frac{\hat{\lambda}_c(\omega)}{2} \right\|_2^2 \right\} \tag{8}$$

Then the sub-optimization model is obtained by replacing

The ω in the front term with $\omega - \omega_k$, and converting the integral to the non-negative frequency domain:

$$\hat{u}_{k,c}^{n+1} = \arg\min_{u_{k,c}} \left\{ \int_0^\infty \left(4\alpha(\omega-\omega_k)^2 \left| \hat{u}_{k,c}(\omega) \right|^2 + 2\left| \hat{x}_c(\omega) - \sum_i \hat{u}_{i,c}(\omega) + \frac{\hat{\lambda}_c(\omega)}{2} \right|^2 \right) d\omega \right\} \tag{9}$$

Obtain the following frequency domain update:

$$\hat{u}_{k,c}^{n+1}(\omega) = \frac{\hat{x}_c(\omega) - \sum_{i \neq k} \hat{u}_{i,c}(\omega) + \frac{\hat{\lambda}_c(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (10)$$

- Update center frequency

Given that neither of the last two terms of the Lagrangian is dependent on ω_k, the related problem can be simplified as follows:

$$\omega_k^{n+1} = \arg\min_{\omega_k} \left\{ \sum_c \left\| \partial_t \left[u_+^{k,c}(t) e^{-j\omega_k t} \right] \right\|_2^2 \right\} \quad (11)$$

According to the Plancherel Theorem, Eq. (11) is transformed into the problem in Fourier domain as follows:

$$\omega_k^{n+1} = \arg\min_{\omega_k} \left\{ \sum_c \int_0^\infty (\omega - \omega_k)^2 \left| \hat{u}_{k,c}^{n+1}(\omega) \right|^2 d\omega \right\} \quad (12)$$

By setting the first derivative of the quadratic function to zero and adding some simple algebraic operations, the sum of the above quadratic functions is minimized, and the following relations are obtained:

$$\omega_k^{n+1} = \frac{\sum_c \int_0^\infty \omega \left| \hat{u}_{k,c}^{n+1}(\omega) \right|^2 d\omega}{\sum_c \int_0^\infty \left| \hat{u}_{k,c}^{n+1}(\omega) \right|^2 d\omega} \quad (13)$$

In the MVMD algorithm, we note that the devotion of power spectrum in C channels is all considered during the whole update of ω_k for every mode.

B. The Proposed Method

Given that practical applications of the original MVMD algorithm need to preset the number of decomposed modes K, we propose a complete set of mechanical fault diagnosis method based upon MVMD. The detailed steps are presented as bellow.

a) *Load multi-channel fault signal data:* Acquire the multichannel vibration signal from the running rotating machines and load them into the MVMD.

b) *Determine the value of parameter K: K* is selected on the basis of the following scheme as shown in Fig. 1. Firstly, for different K, the correlation coefficients of the current intrinsic mode function (IMF)[19] components are calculated by using VMD decomposition. Secondly, the kurtosis of the maximum correlation component under different K values is compared. Lastly, the K value that produces the largest kurtosis is selected as the parameter value.

c) *MVMD is Performed for Decomposition:* Set the initial parameters of the MVMD, run the program, and obtain the decomposition results of input signals.

d) *Select the Optimal Mode:* Compare the kurtosis of each mode component of the decomposed multi-channel signal, and choose the best component with the largest kurtosis in each channel.

e) *Envelope Analysis of the Optimal Mode:* The optimal mode components selected by each channel are analysed by the Hilbert envelope demodulation. If there is an obvious harmonic corresponding at the multiple of a certain frequency in the envelope spectrum, it indicates that this frequency is likely to be the fault characteristic frequency.
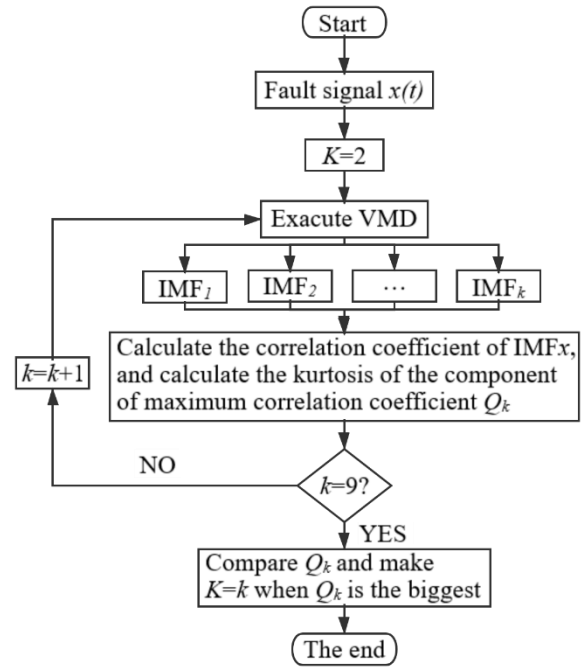
Fig. 1. The method of determining K value

f) *Obtain Results:* Then, compare the extracted value with the theoretical failure frequency calculated according to the failure formula. If the difference is within the allowable error range, it shows that the mechanical fault diagnosis is realized successfully.

III. EXPERIMENT AND VERIFICATION OF MECHANICAL FAULT DIAGNOSIS METHOD BASED ON MVMD

To examine the performance of this new method in actual fault detection, the following faulty bearing data is analyzed in detail in this study. We carried out this experiment on a failure simulator for aero-engine rotor bearing designed and manufactured by Shenyang Aero-engine Design and Research Institute[20]. As shown in Fig. 2, on the housing of the bearing seat, there are three accelerometers mounted to collect the vibration signals with 10 kHz sampling rate. A tachometer was installed to take a measurement the rotating speed of the shaft. An outer ring failure was seeded in the tested bearing with the type of HRB6304, and the outer race failure frequency f_i of this bearing is 64.15 Hz.

Fig. 2. The test rig of the aero-engine rotor-bearing fault simulator
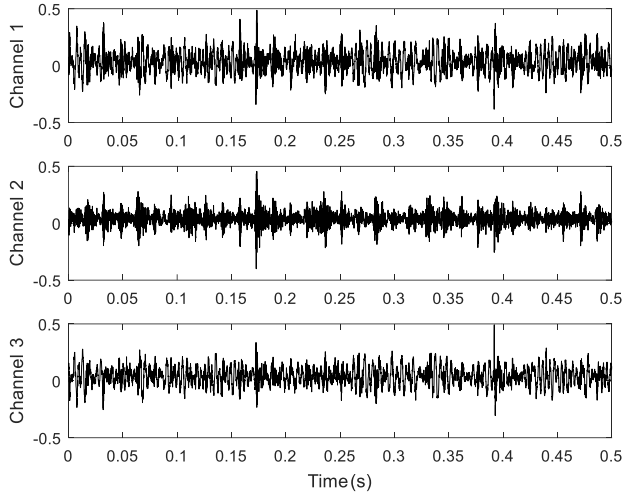
Fig. 3. Three-channel experimental signal
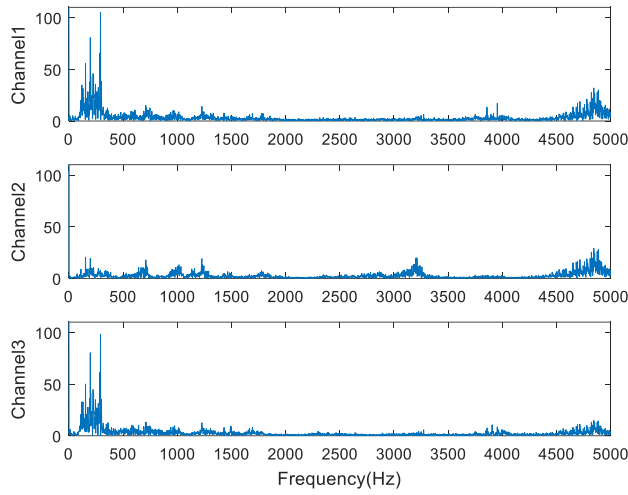
Fig. 4. Spectrum of the three-channel experimental signal

The waveforms of the experimental signal are drawn in Fig. 3 and Fig. 4. As demonstrated in Fig. 5(a), the fault frequency are nearly invisible because of noise interference in the original input signal. Then, this proposed method and the original VMD method are used respectively.

Subsequently, as illustrated in Fig. 5(b), the typical characteristics of outer ring fault are exhibited clearly in the envelope analysis of optimal components, containing the faulty frequency f_i and its harmonics, which conforms with the actual situation. This shows that this new method has a superior noise reduction function and successfully extracts the characteristic frequency of outer ring fault for rolling bearing. Instead, as Fig. 5(c) shows, it is obvious that the original VMD method can only identify the fault characteristic frequency for the channel with less noise. We can also see that the proposed method can recognize and adjust the oscillation of every channel and mode when processing multi-channel signal, so that every channel can extract fault characteristic frequency effectively. All this shows that this new method has a much better performance over the original VMD method in characteristic frequency extraction and background noise filtering.
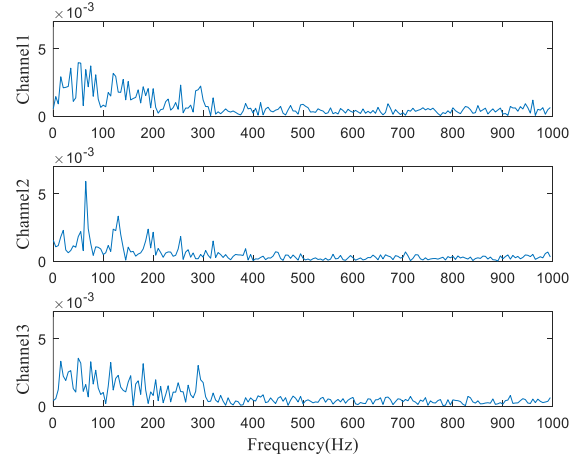
(a) Envelope spectrum of three-channel rolling bearing fault experimental signal
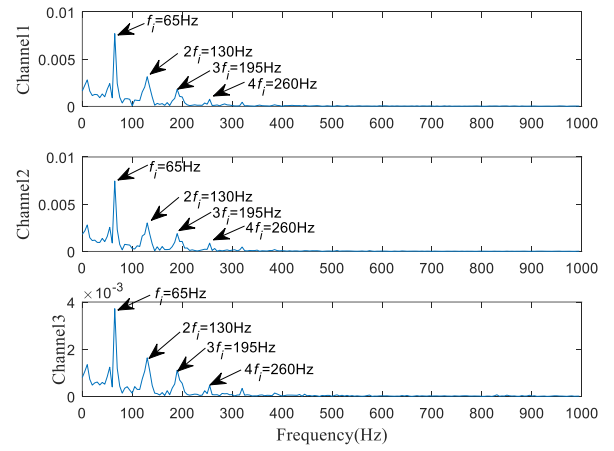
(b) Envelope spectrum obtained by the proposed method
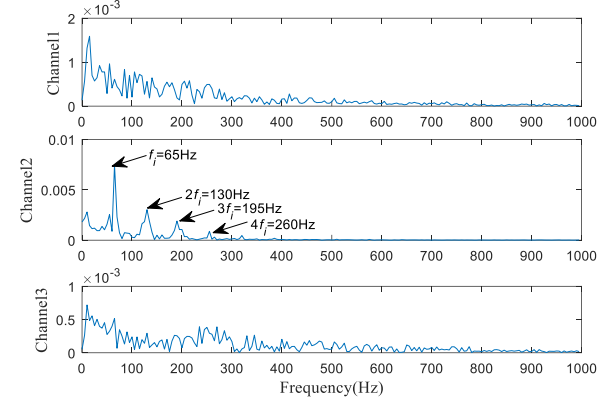
(c) Envelope spectrum obtained by original VMD

Fig. 5. The processing results of three-channel experimental signal at different measuring positions. a) Envelope spectrum of three-channel rolling bearing fault experimental signal; b) Envelope spectrum obtained by the proposed method; c) Envelope spectrum obtained by original VMD

IV.   Conclusion

Aiming at solving the drawbacks of VMD and applying MVMD algorithm to aeronautic industry, a new bearing failure diagnosis method based upon MVMD is put forward for multi-channel signal analysis in this paper. The method of MVMD algorithm is firstly utilized in the field of mechanical failure diagnosis. Moreover, a strategy for selection of parameter K in MVMD is proposed by combining correlation coefficient and kurtosis criterion. Analysis results of this

experimental case confirm not only the superiority of this new method over the original VMD method, but also its availability of improving the performance of fault diagnosis.

REFERENCES

[1] S. Wang, I. W. Selesnick, G. Cai, "Nonconvex Sparse Regularization and Convex Optimization for Bearing Fault Diagnosis," IEEE Transactions on Industrial Electronics, vol. 65, pp. 2727–2736, 2018.

[2] D. Verstraete, E. Droguett and V. Meruane, "Deep semi-supervised generative adversarial fault diagnostics of rolling element bearings," Structural Health Monitoring, vol. 19, pp. 390–411, 2019.

[3] X. Zhao and M. Jia, "A novel unsupervised deep learning network for intelligent fault diagnosis of rotating machinery," Structural Health Monitoring, 2019.

[4] S. Wang, G. Cai and X. Chen, "Matching demodulation transform and synchrosqueezing in time-frequency analysis," IEEE Transactions on Signal Processing, vol. 62, pp. 69–84, 2014.

[5] S. B. Wang, X. F. Chen, G. Cai, B. Q. Chen and X. Li, "Matching demodulation transform and synchrosqueezing in time-frequency analysis," IEEE Transactions on Signal Processing, vol. 62, pp. 69–84, 2014.

[6] C. Li, V. Sanchez, G. Zurita, L. M. Cerrada and D. Cabrera, "Rolling element bearing defect detection using the generalized synchrosqueezing transform guided by time-frequency ridge enhancement," ISA Transactions, vol. 60, pp. 274–284, 2016.

[7] J. Wang, Q. He and F. Kong, "Multiscale envelope manifold for enhanced fault diagnosis of rotating machines," Mechanical Systems & Signal Processing, vol. 52–53, pp. 376–392, 2015.

[8] A. Grossmann and J. Morlet, "Decomposition of hardy functions into square integrable wavelets of constant shape," Siam Journal on Mathematical Analysis, vol. 15, pp. 723–736, 1984.

[9] Z. Feng, D. Zhang and M. Zuo, "Adaptive mode decomposition methods and their applications in signal analysis for machinery fault diagnosis: A review with examples," IEEE Access, vol. 35, pp. 108–126, 2017.

[10] J. Smith, "The local mean decomposition and its application to EEG perception data," Journal of the Royal Society Interface, vol. 2, pp. 443–454, 2005.

[11] N. E. Huang, Z. Shen and S. R. Long, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proceedings Mathematical Physical & Engineering Sciences, vol. 454, pp. 903–995, 1998.

[12] Y. Lei, J. Lin and Z. He, " A review on empirical mode decomposition in fault diagnosis of rotating machinery," Mechanical Systems & Signal Processing, vol. 35, pp. 108–126, 2013.

[13] J. Gilles, "Empirical wavelet transform," IEEE Transactions on Signal Processing, vol. 61, pp. 3999–4010, 2013.

[14] O. Singh and R. K. Sunkaria, "An empirical wavelet transform based approach for multivariate data processing application to cardiovascular physiological signals," Bio-Algorithms and Med-Systems, vol. 14, pp. 903–5, 2018.

[15] D. Zosso and K. Dragomiretskiy, "Variational mode decomposition," IEEE Transactions on Signal Processing, vol. 62, pp. 531–544, 2013.

[16] C. Shen, X. Jiang and Z. Zhu, "Initial center frequency-guided VMD for fault diagnosis of rotating machines," Journal of Sound and Vibration, vol. 435, pp. 36–55, 2018.

[17] N. U. Rehman and H. Aftab, "Multivariate variational mode decomposition," IEEE Transactions on Signal Processing, vol. 67, pp. 6039–6052, 2019.

[18] S. Boyd, N. Parikh and E. Chu, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," Foundations and Trends in Machine Learning, vol. 3, pp. 1–122, 2011.

[19] T. Y. Hou and Z. Shi, "Adaptive data analysis via sparse time-frequency representation," Advances in Adaptive Data Analysis, vol. 3, pp. 1–28, April 2011.

[20] http://ides.nuaa.edu.cn (Accessible by Dec. 02, 2019).

Development and Application of Aero-engine Experimental Data Mining Algorithm Library

Zichen Yan
College of Civil Aviation
Nanjing University of Aeronautics and
Astronautics
Nanjing, China
308148302@qq.com

Haonan Zhang
Technology and Engineering Center for
Space Utilization, Chinese Academy of
Sciences
University of Chinese Academy of
Sciences
Beijing, China
zhanghaonan0122@163.com

Jianzhong Sun*
College of Civil Aviation
Nanjing University of Aeronautics and
Astronautics
Nanjing, China
sunjianzhong@nuaa.edu.cn

Yi Yang
College of Civil Aviation
Nanjing University of Aeronautics and
Astronautics
Nanjing, China
yy1458342850@nuaa.edu.cn

Guangwei Xia
System Engineering Department
AECC Hunan Aviation Powerplant
Research Institute
Zhuzhou, China
1073578348@qq.com

Abstract—**This paper presents the application of several anomaly detection algorithms in experiment data from engine test bed. Several anomaly detection algorithms are programmed in Python language and integrated into an algorithm library named PyPEFD (Python Package for Engine Fault Detection). The algorithm library includes Gaussian Mixture Model, Feature Weighted Fuzzy Compactness and Separation (WFCS), Sequential Probability Ratio Test (SPRT), Variational Autoencoder, Dynamic Time Warping, Mahalanobis Distance, Singular Value Thresholding, Random Forest and Multivariate State Estimation Technique. These algorithms can analyze the structure and characteristics of the engine test data, and mine the hidden fault information in the data, so as to detect the fault or fault trend of aero-engine test data. This paper also presents a preview of the algorithm library.**

Keywords—aero-engine test data, anomaly detection, python algorithm package, data analysis

I. INTRODUCTION

A. Background

With the with the modernization and advancement in data acquisition techniques, and the adoption of various new sensors (such as gas path electrostatic sensor [1], blade tip timing and gap sensor [2], oil path abrasive particle sensor [3], etc.), the breadth and depth of engine condition monitoring has been expanded. The fault detection based on condition monitoring techniques and machine learning have tremendous potential. It provides a rich structured data source for improving and perfecting the engine test evaluation system. How to mine and integrate these multi-source, multi-modal, multi-temporal and spatial scale "big data" so as to improve the utilization of test data and the engine test evaluation system, is the frontier issue of multi-modal data analysis and research in the field of engine test [4].

Throughout the research and development process of aeroengine data analysis and condition monitoring, it can be divided into three stages [5]: Manual judgments stage, Sensors monitoring stage and artificial intelligence stage.

At present, the research and development of anomaly detection algorithms in foreign countries is more advanced. William R. Jacobs, Huw L. Edwards and others proposed a fault detection method: using Gaussian mixture model to analyze target data, and then evaluate the posterior probability by Hidden Markov Model to determine the system's status at each data point [6]; Consumi M et al. used Bayesian inference method to monitor the state of turbojet engines [7]. Dewallef P et al. use an improved Kalman filter model to perform online performance detection and fault diagnosis for engines [8]; Seo D H et al. proposed a support vector machine and neural network fusion method for engine state monitoring [9]. Puranik et al. used neighborhood-based methods such as K-nearest neighborhood (KNN) and local outlier factor (LOF) to perform in-depth quantitative analysis of the clustering results of flight data [10]. In terms of detecting instantaneous abnormal points, Manukyan A et al. also proposed an online data anomaly detection method based on KNN [11].

However, compared with the advanced technology and manufacturing experience of European and American countries, many detection technologies in china are still in research stage, and the self-developed algorithms need to be improved. At present, manual analysis or foreign commercial software are still used. This paper presents an algorithm library mainly for anomaly detection of aero-engine, which is not frequently researched in fault detection area.

B. Data Introduction
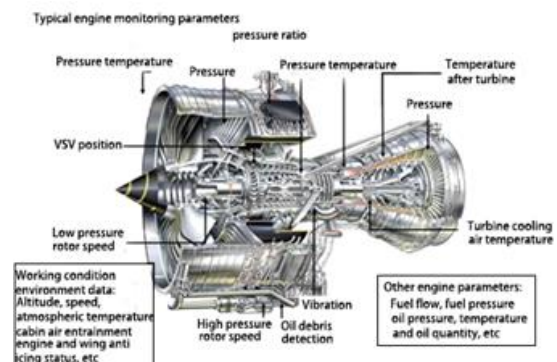
Fig. 1. Aero-engine typical monitoring parameters.

Figure 1 shows the online state monitoring parameters of

---

a typical modern turbofan engine, covering the main components of the engine gas circuit and important accessory systems (accessory systems such as lubricating oil and fuel control), etc. The frequency of the airborne data acquisition system is roughly between 0.25 and 4 Hz. The entire experiment has collected a large amount of test data for further analysis. The new generation of aero-engines is equipped with more advanced and specialized status data collection and processing systems (Engine Monitoring Unite, EMU) (independent of EEC), which expands the scope and depth of status data collection. What's more, new health monitoring sensors are adopted, such as Tip timing sensors, gas path static sensors, oil path wear particles sensors, etc., which enrich the information source of engine health evaluation.

In this paper, the fault samples are collected from the fault data of the engine test bed. The test bed can simulate a variety of engine failures and reflect them in the form of data sets. The purpose of this paper is to use some anomaly detection algorithm programs to identify the faults in the test bed data, and to compare the differences of the results.

II. ANOMALY DETECTION ALGORITHMS

A. Gaussian Mixture Model, GMM

The Gaussian Mixture Model is a powerful tool for classification problems. It is a clustering algorithm based on normal distribution models. The essence of this method is Gaussian fitting, combined with Expectation-Maximization (EM) algorithm for parameter estimation [11]. Generally, the Gaussian mixture model is the expansion of the single Gaussian model, using a combination of multiple Gaussian distributions to describe the data distribution, so as to reach the purpose of clustering [12].

In the anomaly detection of engine data, Gaussian mixture model is suitable for distinguishing the fault data, because when the data is abnormal, the parameters' mean and variance will also change. These changes will be identified by the algorithm and divided into new cluster centers [13]. In some of the cases, GMM are used along with other algorithms for an efficient fault detection system.

B. Mahalanobis Distance, MD

Mahalanobis Distance is a distance concept proposed by Mahalanobis. This distance concept can offset the influence of different scales of each dimension parameter, and the result will be more reliable [14].

The covariance matrix can measure the correlation between data in various dimensions [15]. In the actual fault detection case, the mahalanobis distance list, which contains the mahalanobis distance between each data point and the standard point is used to represents the entire state change process. The points with larger distance indicate that they have a greater deviation from the normal state. These are the abnormal points that shall be analysis.

C. Feature Weighted Fuzzy Compactness and Separation, WFCS

WFCS is an improved clustering method based on fuzzy clustering, whose full name is Feature Weighted Fuzzy Compactness and Separation (WFCS). This clustering algorithm pays more attention to the characteristics of the data in each dimension, and iteratively optimizes the

clustering center through a weighted method, and has a high clustering accuracy.

The application method of WFCS in engine data anomaly detection is similar to the Gaussian mixture model, which is used to distinguish between normal data and abnormal data.

D. Random Forest, RF

The random forest is among the most practical learning methods used in machine learning. It is composed of several randomly generated decision trees, among which the decision trees are not related to each other [16]. The essence of the method is to classify or regress samples [17].

In the actual engine fault detection, the normal labeled data (usually refers to the data without fault) can be trained to get the model. The predicted value is obtained through the model, and then the deviation between the predicted value and the real value can reflect the fault signal of the sample, so as to achieve the purpose of fault detection.

E. Multiple State Estimation Technique, MSET

MSET have become renowned in fault detection system due to their reliability, adaptability and robustness. The model is trained by memory matrix with normal data, and according to the weight parameter, the new observation value is estimated, so as to judge the state of the system [19]. The model uses the new input vector and all normal observation vectors in the memory matrix to obtain the predicted value [20].

III. INTRODUCTION OF ALGORITHM LIBRARY

Firstly, the algorithm programs are written and made sure that they can be successfully called. Then, the algorithm programs are packaged as PyPEFD algorithm library (including random forest, multivariate state estimation, Gaussian mixture, etc.) by using the setup function in the setup tools module.

By downloading and installing the PyPEFD python package, algorithms can be directly imported.

Fig. 2. Installation window of PyPEFD package.

Figure 3 shows all 10 algorithms in the library (all abbreviated). After the installation, the command: "*from PyPEFD.models Import...*" is used to import the algorithms. The detailed introduction of the algorithms package is in table 1.

Table 1 introduces the types of algorithms included in the algorithm library and their application in engine abnormality detection. Because the actual engine failure situation is very complicated, the diversity of algorithms needs to be guaranteed in order to improve detection efficiency. These algorithms partly come from previous projects, which have achieved good application results.
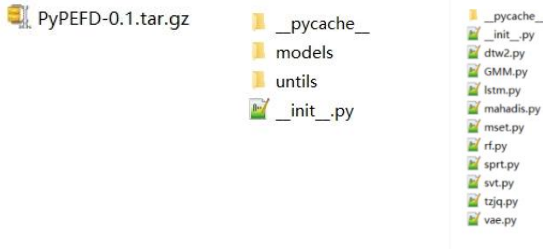
Fig. 3. PyPEFD content.

TABLE I.   TABLE TYPE STYLES

Algorithm	Description
dtw2 (Dynamic Time Warping, DTW)	The DTW algorithm is often used to compare the similarity of time series. After inputting the sample, the algorithm will calculate and output a series of "minimum regularization distance" that characterizes the sequence similarity. The higher the similarity, the smaller the minimum regularization distance.
GMM (Gaussian Mixture Model)	The GMM can represent an arbitrarily complex distribution given a sufficient quantity of Gaussian components. After the data is input, the algorithm can output a list of labels that characterize the clustering results (that is, which category each data point belongs to) and a visualization graph of the clustering results.
LSTM (Long Short-Term Memory)	Long-term and short-term memory network is a kind of time recurrent neural network, which is suitable for dealing with time series problems. Here it is used for sequence prediction. By inputting the known series of parameters, and then it can predict the unknown series through the training model.
Mahadis (Mahalanobis Distance)	Before calculation, a standard point is required and should be placed in the first line of the sample. Then the algorithm will calculate the mahalanobis distance between each data point and the standard point, and output a list for users to observe deviation from the standard situation.
Mset (Multiple State Estimation Technique)	MSET uses normal data to build training matrix, and the input is a new observation matrix. After calculation, the MSET outputs the estimation matrix. The state of the system can be judged by estimating the deviation between the matrix and the observation matrix.
Rf (Random Forest)	Random forest is a combination of multiple decision trees for the decision making. It has multitude of decision trees at a training time. It outputs the mode of the class of the individual tree.
Sprt (Sequential Probability Ratio Test)	SPRT is an unsupervised inspection method that determines whether to continue operation based on existing results. It is mainly used to check whether the sequence structure changes abnormally, and is mostly used for time series sequences.
Svt (Singular Value Threshold)	SVT is used for matrix completion, input sparse matrix, and then output reconstruction matrix.
Tzjq (Feature Weighted Fuzzy Compactness and Separation，WFCS)	WFCS clustering is an improved clustering method based on traditional fuzzy clustering. After the sample is input, the algorithm begins to iterate, continuously updating the cluster centers of each category, and finally outputs a list of labels that characterize the clustering results.
Vae (variational autoencoder)	The learning goal of the network is to make the distribution function of variables approximate the real distribution function. After training the model, the output result is used to judge whether there is a fault.

IV. CASE STUDY

A. Oil System Fault Detection

This subsection presents a case study of fault detection in a sudden drop in oil pressure. According to the influence of the fault, some parameters that can reflect the drop in lubricating oil pressure in the fault data set are screened out as input parameters of the algorithm.

The parameter "XX" is collected from the engine test bed which reflects the actual change of the oil pressure difference in the engine, as shown in the figure 4.
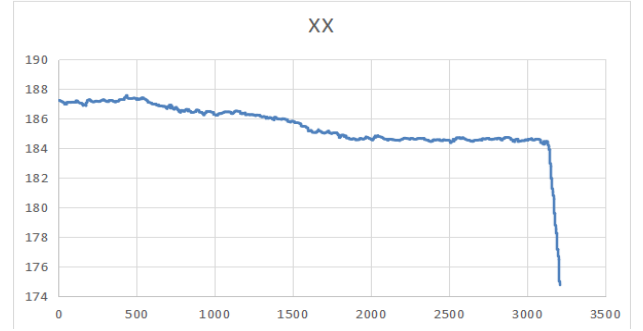
Fig. 4. Change of oil pressure difference.

It should be noted that the last 3000th data point dropped sharply afterwards, that is because the engine was shut down at that time, not a manifestation of the fault.

The figure 5 shows a partial display of the fault data set. The fault data set contains a large number of parameters, such as pressure, air flow and temperature reflecting the performance of the gas circuit, or control parameters reflecting the pilot's operation process. Each column contains a parameter's sequence in time, and the simulation is conducted at a frequency of 50 Hz.

Fig. 5. Fault data set.

In this case, GMM, WFCS, RF, and MSET are used to detect failure trends in lubricating oil fault data. The following are the operating results of each algorithm.
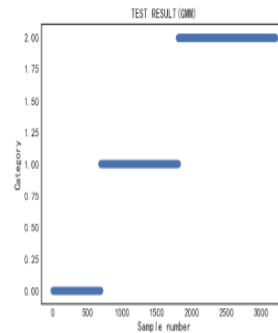
Fig. 6. Test result (GMM)   Fig. 7. Test result (WFCS)
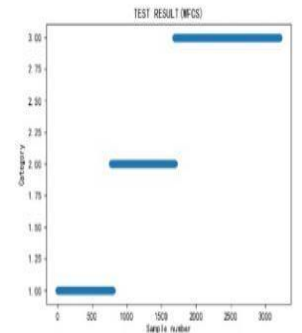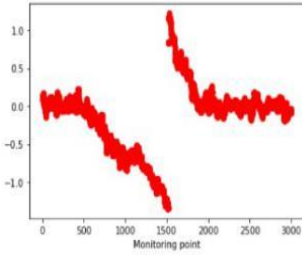
Fig. 8. Test result (RF)
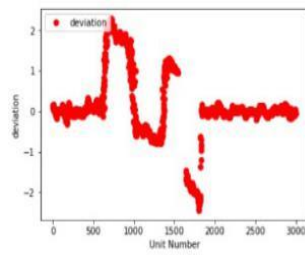
Fig. 9. Test result (MSET)

1) Experimental Analysis

The oil pressure drop is a continuous fault, and the purpose of this case is to detect the fault area where the oil pressure changes. Then there are two ways to accomplish this goal: one is to start from the change of the parameters' mean and variance, and the other is using deep learning algorithms to calculate deviation values.

Generally speaking, the cluster type algorithm is very sensitive to changes to the parameters' mean and variance, so two clustering algorithms, GMM and WFCS, are chosen for detection. The number of categories is set to 3, that is, the entire fault data set is divided into three stages, which represent the stages before, during and after the drop in the oil pressure difference.

Figure 6 and Figure 7 show the detection results of the clustering algorithms. Both algorithms successfully divide the data points into three categories. Here we focus on category 1 in Figure 6 and category 2 in Figure 7, because this is a pressure dropping state between the initial state and the end state, which is considered as a fault state.

The results need to be compared with the parameter XX. Figure 4 shows the actual change of the oil pressure difference, which roughly started to decrease from the 500th data point until the 1900th data point. The whole decline process lasted 29 seconds, totally 1400 data points. The test result shows that GMM correctly detected 1,115 points, while WFCS correctly detected 1,008 points. The result is similar to the real situation, and the accuracy of the GMM algorithm is higher.

RF and MSET are two machine learning algorithms. They can build connections between parameters of different dimensions, and through these connections the algorithms can predict a specific parameter based on the remaining parameters. Here the data with no change in oil pressure is used as the training set. The training set is input into the algorithms so as to predict the parameter "XX" in normal condition. Then the deviation can be calculated.

Figure 9 and figure 10 show the deviation between the actual value of the parameter XX and the predicted "XX" of the algorithms. It can be seen that the deviation values of the oil pressure difference show a certain abnormal trend when the oil pressure difference decreases, which indicates that the data segment between 500 and 2000 has become abnormal.

B. SurgeFault

1) Wind-Driven Surge Fault Detection

This subsection presents a case study of fault detection in the wind-driven surge, and uses the anomaly detection algorithm to detect the abnormality of the data sequence.

The parameter NN is a signal from engine test bed that characterizes whether surge occurs, and the increase in the value of NN means that surge happens. The purpose of the case in this section is to compare the detection results of different algorithms with the actual situation of the NN, so as to judge the accuracy of the fault recognition. The results are shown below:
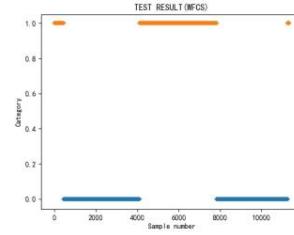
(a) Test result (WFCS)
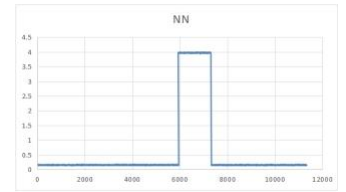
(b) Test result (engine test bed)

Fig. 10. WFCS test result (sample 1)
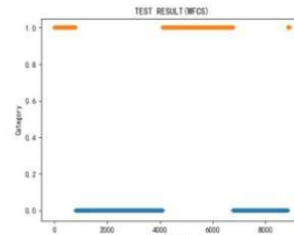
(a) Test result (WFCS)
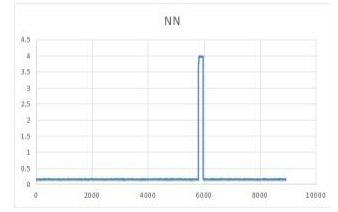
(b) Test result (engine test bed)

Fig. 11. WFCS test result (sample 2)
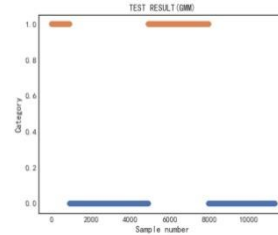
(a) Test result (GMM)
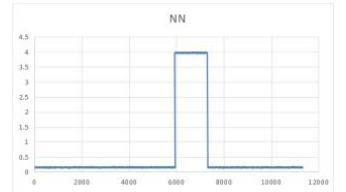
(b) Test result (engine test bed)

Fig. 12. GMM test result (sample 1)
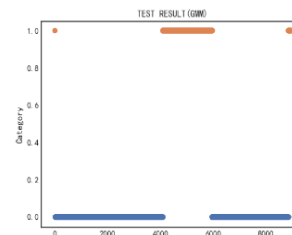
(a) Test result (GMM)
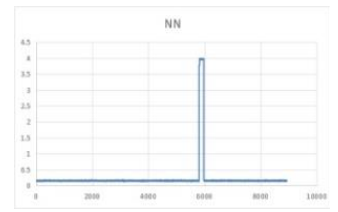
(b) Test result (engine test bed)

Fig. 13. GMM test result (sample 2)
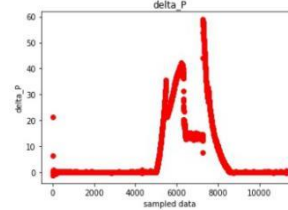
Fig. 14. Deviation value of pressure predicted by RF model
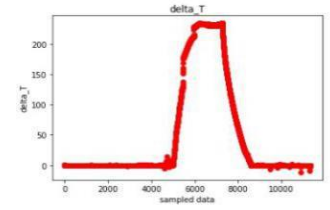
Fig. 15. Deviation value of temperature predicted by RF model
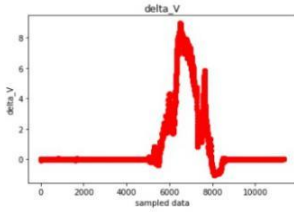
Fig. 16. Deviation value of RF model for vibration prediction
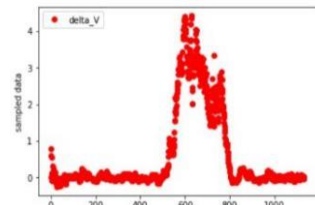
Fig. 17. Deviation value of vibration predicted by MSET model
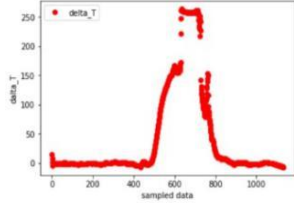
Fig. 18. Deviation of temperature predicted by MSET model
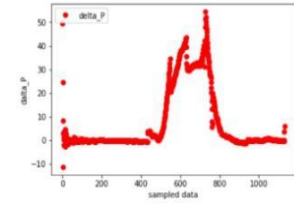
Fig. 19. Deviation of pressure predicted by MSET model

2) Forced Surge Failure

Forced surge usually refers to artificially causing the engine to surge, which is used to verify the aerodynamic stability of the engine. In order not to damage the structure of the engine, the surge time shouldn't be too long. In the forced surge fault case, MD, RF and MSET are chosen to complete the detection. Before the detection, the proper parameters should be chosen so as to make results more precise.

A "Pj1" parameter is provided in the sample to show the location of forced surge, which provides a reference for the detection results of the algorithms. The comparison of the detection results and the parameter "Pj1" is shown below.
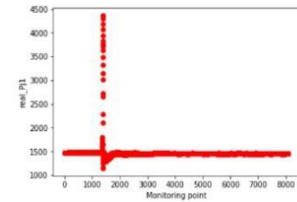
Fig. 20. Pj1 parameters
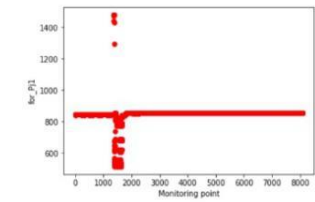
Fig. 21. Prediction of Pj1 results by RF model
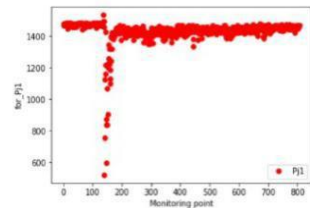
Fig. 22. Prediction of Pj1 by MSET
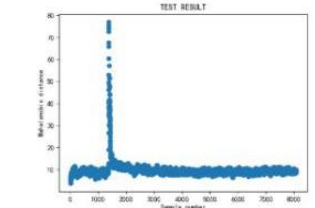
Fig. 23. Test result (MD)

3) Experimental Analysis

Surge fault diagnosis in an aircraft gas turbine engine involves identification of the fault type, and location and quantification of the fault level. Surge will cause significant changes in some parameters before it occurs, such as pressure drop and rotating speed fluctuations, etc. The algorithm must discover the abnormal trend of these parameters in advance. In wind-driven surge detection, the goal has been accomplished very well.

It is observed from figure 10 and 11 that samples of surge fault are efficiently classified. In Figure 10(b), the surge in sample No. 1 occurred at the 6000th sample point

and lasted to about 7000th, while the WFCS algorithm detected anomalies at the 5000th sample point, and the subsequent 2000 sample points are classified as fault types, which is earlier than the NN parameter. The second sample is even more obvious. The WFCS algorithm detected anomalies at the 4000th sample point, which is much earlier than the parameter NN.

Figures 12 and 13 show the detection results of Gaussian mixture model clustering, which are very similar to the results of the WFCS.

Next is the detection results of RF and MSET. The algorithms use the experimental data to predict and calculate the deviation value with the real pressure, temperature and vibration parameters (these parameters are influenced by surge), then compare it with the surge parameter "NN" provided by the engine test bed.

Figures 14, 15 and 16 show the deviation values of pressure, temperature and vibration predicted by RF model. It can be seen that the deviation value increases significantly when surge occurs, indicating that surge occurs. According to the deviation value in the figure, it even changes ahead of the surge parameter NN.

Figure 17,18,19 shows the prediction results of MSET (samples are reduced at equal intervals for training purposes), which is very similar to the detection results of RF. The deviation values of three parameters at the position where surge occurs rise sharply, and the abnormal position is slightly ahead of the parameter "NN", which means the abnormal situation is successfully detected before the surge occurs.

In forced surge case, the situation is different. Due to short duration of the forced surge, the detection effect of the clustering algorithms is unsatisfactory, so the goal has changed from preventing the occurrence of surge to finding the specific location of surge. Mahalanobis Distance (MD) is used in this case. It can composite the abnormal trend of each dimension parameter to calculate a distance describing the deviation degree. Figure 23 is the result of the MD algorithm. It can be seen that the value of mahalanobis distance rise sharply in the same position where the parameter "Pjl" rise, which means the algorithm finds the correct position. The RF and MSET algorithms also accomplish the task, and the detection method is the same as above.

In summary, unsupervised learning algorithms (clustering algorithms, Mahalanobis Distance, etc.) are more suitable for real-time monitoring, because they can quickly get the test results and reflect the changes of parameters in real time, helping pilots or engineers find problems in time. However, their disadvantage is lower accuracy. In contrast, supervised learning algorithms have higher accuracy, but due to their long training time and high requirements for training data, they are more suitable for offline detection, providing more accurate results.

V. CONCLUSION

This paper has presented a detailed review of a python - based algorithm library for engine fault detection, including a variety of anomaly detection algorithms (Mahalanobis Distance, Gaussian mixture model, WFCS, SPRT, MSET, random forest, VAE, etc.). These methods can identify the

health status of the engine data, which is of great significance to the line maintenance of aero-engine.

The innovations of the algorithm library are listed below:

1) This algorithm library is specifically established for engine fault data detection.

2) There are various of algorithms and comprehensive functions in the library, which can make an analysis from different angles.

3) This algorithm library has been integrated into a python module, which is convenient for users to install and apply.

In the case study part of the paper, the anomaly detection algorithms like GMM, WFCS, RF, MSET and MD have been discussed thoroughly, and the detection is performed based on the actual test data. The results show that the algorithm can effectively detect a variety of faults, including oil system faults and surge faults.

The successful application of these algorithms proves the reliability and efficiency of the algorithm library, and establishes a foundation for our follow-up research on other types of fault detection. Although the method presented in the paper can perform in real time, various operating conditions still need to be investigated. The future research should focus on improving the accuracy of detection and the installation of new algorithms.

REFERENCES

[1] Valentina Zaccaria, Mikael Stenfelt. Fleet. Monitoring and Diagnostics Framework Based on Digital Twin of Aero-Engines//Proceedings of ASME Turbo Expo 2018 Turbomachinery Technical Conference and Exposition, June.2018:GT2018-76414.

[2] Ho L M. Application of Adaptive Thresholds in Robust Fault Detection of an Electro-Mechanical Single-Wheel Steering Actuator. IFAC Proceedings Volumes, 2012, 45(20):259-264.

[3] Palacios A, Martníez A, Sánchez L, et al. Sequential pattern mining applied to aeroengine condition monitoring with uncertain health data. Engineering Applications of Artificial Intelligence, 2015, 44: 10-24.

[4] Frolik J, Abdelrahman M, Kandasamy P. A confidence-based approach to the self-validation, fusion and reconstruction of quasi-redundant sensor data. Instrumentation and Measurement, IEEE Transactions on, 2001, 50(6): 1761-1769.

[5] Ying Y, Cao Y, Li S, et al. Nonlinear Steady-State Model Based Gas Turbine Health Status Estimation Approach with Improved Particle Swarm Optimization Algorithm. Mathematical Problems in Engineering, 2015.

[6] William R. Jacobs, Huw L. Edwards. Gas Turbine Engine Condition Monitoring Using Gaussian Mixture and Hidden Markov Models. International Journal of Prognostics and Health Management, 2018.

[7] Consumi M, d'Agostino L. Monitoring and fault diagnosis of a turbojet by Bayesian inference//13th International Symposium on Air Breathing Engines (ISABE 97-7146), September. 1997: 1082-1096.

[8] Dewallef P, Léonard O. On-line performance monitoring and engine diagnostic using robust Kalman filtering techniques. ASME Paper No. GT2003-38379, 2003.

[9] Seo D H, Roh T S, Choi D W. Defect diagnostics of gas turbine engine using hybrid SVM-ANN with module system in off-design condition. Journal of mechanical science and technology, 2009, 23(3): 677-685.

[10] Breunig M M, Kriegel H P, NG R T, et al. LOF: identifying density-based local outliers. ACMSIGMOD International Conference on Management of Data, 2000: 93-104.

[11] Manukyan A, Olivares-Mendez M A, Voos H, et al. Real time degradation identification of UAV using machine learning techniques. IEEE International Conference on Unmanned Aircraft Systems, 2017: 1223-1230.

[12] S. Yeom. Infrared Image Segmentation Based on Region of Interest extraction with Gaussian Mixture Modeling. International Society for Optics and Photonics, 2017, 10202:102020C.

[13] WILSON R. Multiresolution Gaussian Mixture Models: Theory and Application. Research Report RR404, Deparhment of Computer Science, University of Warwick, UK, 1999: 1-10.

[14] S.Kumar,V Sotiris, M.Pecht. Mahalanobis Distance and Projection Pursuit Analysis for Health Assessment of Electronic Systems. IEEE Aerospace conference,2008,1640(2):1-9.

[15] Cai Jingying, Xie Fuding, Zhang Yong. Fuzzy c-means algorithm based on adaptive Mahalanobis distances[J]. Computer Engineering and Applications, 2010, 46(34): 174-176.

[16] W. Yan. Application of Random Forest to Aircraft Engine Fault Diagnosis. The Proceedings of the Multiconference on "Computational Engineering in Systems Applications", Beijing, 2006, pp. 468-475, doi: 10.1109/CESA.2006.4281698.

[17] Y. Santur, M. Karaköse, E. Akin. Random Forest Based Diagnosis Approach for Rail Fault Inspection in Railways. 2016 National Conference on Electrical, Electronics and Biomedical Engineering (ELECO), Bursa, 2016, pp.745-750.

[18] L. Breiman. Random forests. Machine Learning, vol. 45, pp. 5-32, 2001.

[19] Peng, J., Xiao, W. D., & Huang, X. P. A Health Monitoring Method Based on Multivariate State Estimation Technique. Applied Mechanics and Materials, 2013, 281, 80–85.

[20] C. Li, J. Sun, H. Zuo, J. Chen. Fault Detection for Air Conditioning System of Civil Aircraft Based on Multivariate State Estimation Technique. 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Shanghai, 2017, pp. 180-185, doi: 10.1109/SDPC.2017.42.

Differentiable Architecture Search for Aeroengine Bevel Gear Fault Diagnosis

Zheng Zhou
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
nauwu2@stu.xjtu.edu.cn

Tianfu Li
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
litianfu@stu.xjtu.edu.cn

Zhibin Zhao
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
zhibinzhao1993@gmail.cn

Chuang Sun
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
ch.sun@xjtu.edu.cn

Ruqiang Yan*
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
yanruqiang@xjtu.edu.cn

Xuefeng Chen
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
chexf@xjtu.edu.cn

Abstract—The deep learning models have swept the field of intelligent fault diagnosis, especially the deep convolutional neural network (DCNN) with ability to automatically extract features from large-scale data. On one hand, as most of the existing DCNN models are specifically designed for image task, direct application of these models to machine fault diagnosis will cause domain bias. On the other hand, it is time-consuming and often dependent on expert knowledge to design a DCNN model for a given fault diagnosis task from a specific dataset. In this paper, a differentiable architecture search method through gradient optimization is proposed to design a new network model for aeroengine bevel gear fault diagnosis. This method can design efficient network from a complex search space for a specific dataset via constructing the hyper-network and gradient-based search strategy. The proxy model and optimization approximation are used to speed up the searching process. Compared with the traditional discrete neural architecture search methods, this method has great superiority in computational cost. The aeroengine bevel gear dataset is used to verify the performance of this method, and only 3 GPU hours are required for this dataset to search the optimized network. The proposed method achieves the state-of-the-art performance under different signal-to-noise ratios via comparisons with other mainstream manually designed networks.

Keywords—fault diagnosis, deep convolutional neural network, differentiable architecture search

I. INTRODUCTION

The health state of the major equipment is highly concerned in modern industry, as its sudden shutdown or failure will bring serious economic losses and safety risks. Prognostic and Health Management (PHM) provides proactive decision making capability for maintenance by monitoring the operating condition of the equipment so as to reduce the maintenance cost. Fault diagnosis, as one important ingredient of PHM, is to identify the fault types and qualitatively determine the health states of equipment. It is located in the upstream of the whole PHM process to provide information for the subsequent prognosis and maintenance decision.

With the development of intelligent sensing and artificial intelligence technology, machine learning (ML) algorithms have been widely used in fault diagnosis as the baseline, such as support vector machine [1] and tree-based methods [2]. However, since the feature extraction of the ML-based fault diagnosis is separated from the classifier, its performance is heavily dependent on the manual features.

Deep learning (DL), especially deep convolutional neural network (DCNN), combines the feature engineering and the classification as an end-to-end optimization pipeline, to realize the automatic feature extraction from large-scale data. Therefore, the overall structure of DCNN can be divided into the deep feature extraction module and the shallow classifier module. The deep module extracts high-quality features from the input data through the convolutional layers, pooling layers, batch normalization layers, etc. Contrast with the shallow classifier module, the deep module affects the feature quality and is more worth studying. There has been several works on this perspective. Li et al. [3] proposed an adaptive channel weighted convolutional layer to rank the information importance of multi-source sensors. Hu et al. [4] proposed an improved deep module to extract multiscale features for the gearbox fault diagnosis. In fact, the performance of the deep module is mainly determined by the parameters in its architecture. This means that the feature engineering paradigm of ML is shifted to architecture engineering in DL.

The DCNN firstly made a breakthrough in the field of computer vision (CV) and has been extensively used as the backbone. For example, the popular high-performance network architectures, such as VGG [5], ResNet [6] and Inception [7], have been adopted to the image data specifically. In fact, the architecture parameters of networks are complex, discrete and disordered hyper-parameters, required to be adjusted from multiple dimensions, such as depth, width and skip connection, etc. Therefore, the trial-and-error experiments are required to design corresponding network architectures for specific tasks, which is time-consuming and relies on expert knowledge.

Additionally, the task of CV is to process the image with rich spatial information, while the task of fault diagnosis is to deal with time sequence signal. The direct application of networks from CV to fault diagnosis is a model migration process, which causes the domain bias.

Neural architecture search (NAS) is a competitive approach for the aforementioned problem. It is an automated technique to design neural network for a specific dataset, and has achieved the state-of-the-art performance in image classification [8], object detection [9] and video understanding [10]. In addition to the CV tasks, NAS is also applicable to natural language processing [11] and speech recognition [12]. As a powerful automation tool for designing networks, NAS is bound to attract attention in various fields.

Scholars engaged in PHM have gradually applied the NAS methods to fault diagnosis and prognosis. Wang et al. [13]

---

*Ruqiang Yan is the corresponding author. (e-mail: yanruqiang@xjtu.edu.cn)

proposed a reinforcement learning-based NAS method for bearing fault diagnosis. Li et al. [14] used weight sharing strategy to speed up the searching process involved in NAS for gearbox fault diagnosis. Li and He [15] proposed a NAS method based on Bayesian optimization for remaining useful life prediction.

According to the optimization methods, inchoate NAS can be divided into two groups: reinforcement learning-based [16] and evolutionary algorithm-based [17]. Both of them suffer from huge amount of computing resource consumption. Recent NAS research has focused on speeding up the searching and evaluation processes, such as weight sharing [18], proxy task [19], performance prediction [20], etc. However, all these methods involve an inevitable procedure that a large amount of candidate networks need to be trained from scratch.

In this paper, we introduce an one-shot NAS method based on gradient optimization, called differentiable architecture search (DARTS) [21]. This method relaxes the search space by mixing the operators to construct the hyper-network, and then the gradient optimization can be used. As the hyper-network contains all the candidate networks, this method avoids the cost of training the large amount of candidate networks from scratch. Therefore, DARTS is more efficient and resource-friendly than the previously nondifferentiable methods. The model proxy strategy and optimization approximation are used to speed up the searching process simultaneously.

The rest of this paper is organized as follows. Section 2 describes the search space, optimization method and speed up strategy of DATRS in detail. In section 3, we perform experimental verification on the multilevel signal to noise ratio (SNR) datasets. Finally, Section 4 summarizes the main conclusions of this paper.

II. METHODOLOGY

The proposed approach for aeroengine bevel gear fault diagnosis is described in Fig. 1. It consists of four steps: data acquisition and preprocessing, searching process, training the sub-network from scratch, and testing. The aeroengine bevel gear fault data in the testing bench are collected, and divided into train set, validation set and test set first. Then DARTS is used to search a customized network for this dataset. The train set is used to optimize the model weights and the validation set is used to optimize the architecture weights. After that, the customized network is trained from scratch based on the train set and validation set. Finally, the aeroengine bevel gear fault diagnosis is inferred based on the trained model.

In this section, we focus on the optimization method of DARTS and briefly introduce its search space and speed up strategy.

A. Search Space

The search space represents the domain priori, which requires a manual assignment, and the operators in the search space have been designed and validated by the professional researchers in time and effort. Besides, the consistent search space ensures the comparability of various algorithms, which is conducive to the algorithm development. For different tasks, it is feasible to design the corresponding search space for NAS, and it is worth noting that the size and complexity of the search space determine the diversity of sub-networks and the computational cost of the searching process.

Unlike simple parameters such as the number of filters or the number of layers, DARTS can design networks from a more complex search space, while ensuring sufficient candidates. The operators in the search space of DARTS are described in Table I.
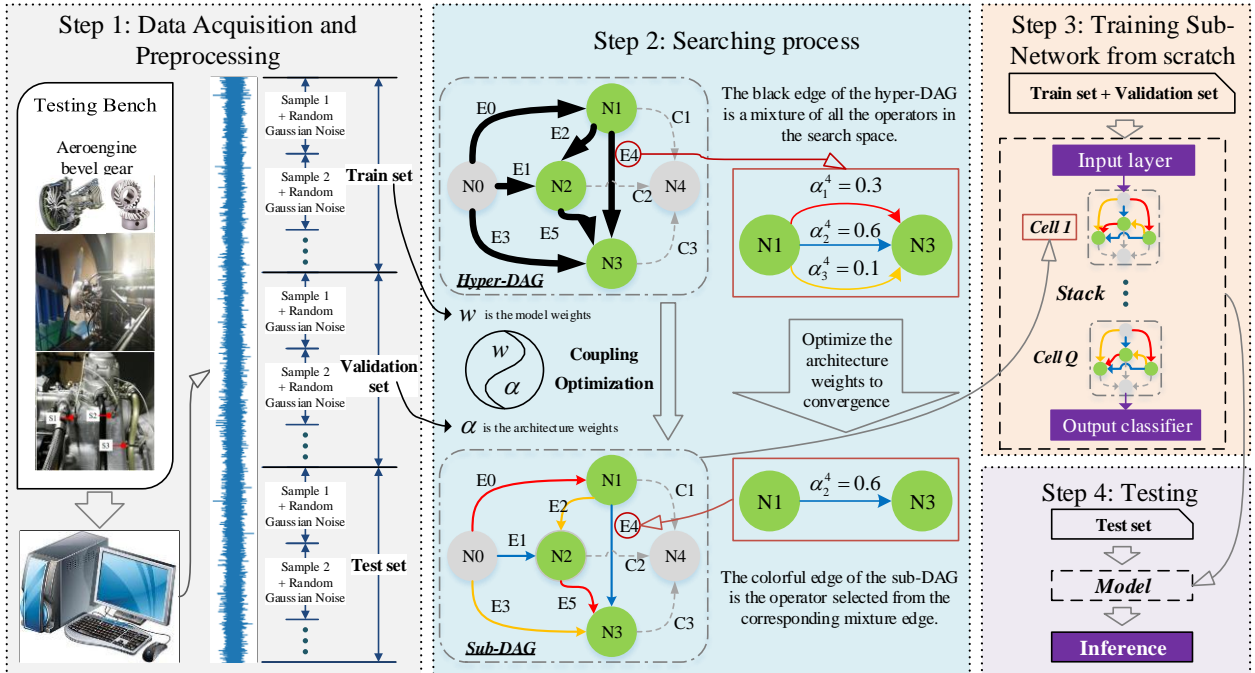
Fig. 1. Proposed NAS-based intelligent fault diagnosis framework for aeroengine bevel gear

TABLE I. THE OPERATORS IN THE SEARCH SPACE OF DARTS

Operator Name	Description
None	Zero
Max_pool	Max pooling
Avg_pool	Average pooling
Skip_connect	Identity mapping
Sep_conv_3	Kernel size =3 separable convolutions
Sep_conv_5	Kernel size =5 separable convolutions
Dil_conv_3	Kernel size =3 dilated separable convolutions
Dil_conv_5	Kernel size =5 dilated separable convolutions

B. Gradient-based Optimization Method

Instead of searching the global network directly, DARTS adopts the proxy model strategy searching only two types of periodic units, that is, normal cell and reduce cell. For instance, the normal cell and reduce cell are shown in the Step 2 of Fig. 1, and their structures are both directed acyclic graph (DAG), which are composed of several nodes and edges. The difference is that the output dimension of normal cell is consistent with the input, while the number of channels in the output dimension of reduce cell is doubled and the length is halved. Therefore, there is no essential difference between normal cell and reduce cell in design. By changing the number of stacks, DARTS can obtain global networks of varying complexity.

The first procedure is to construct a hyper-DAG, which contains all the candidate sub-DAGs. The principle of the hyper-DAG construction is that the input to the intermediate node is the output of all the previous nodes. For instance, as shown in the Step 2 of Fig. 1, the cell has 5 nodes and 9 edges, where {N0, N4} are the input and output nodes of the current cell, {N1, N2, N3} are the intermediate nodes, {E0, E2, E3, E4, E5} are the mixed operators edges, and {C1, C2, C3} are the connecting edges. The nodes represent different feature maps and the edges represent corresponding operations. The input node is a copy of the output node of the previous cell. The intermediate nodes represent the feature maps after complex operations in the search space. The connecting edges splice the feature maps of all the intermediate nodes into the output node of the current cell. The key point of the hyper-DAG is the construction of the mixed operators edges, each of which is superimposed by all the operators in the search space multiplied by a coefficient vector called architecture weight, denoted as α.

The forward calculation of each intermediate node in the cell can be formulated as:

$$f_n(x) = \sum_{i=0}^{n-1} \sum_{m=0}^{M} \alpha_m^{\frac{n(n-1)}{2}+i} OP_m^{\frac{n(n-1)}{2}+i}(f_i(x)), \ n \in \{1,2,3\} \quad (1)$$

where n denotes the order of the intermediate node, M is the size of the search space, and OP represents the corresponding operator.

According to the design criteria of the hyper-DAG, the intermediate nodes of the cell are not independent, as the later nodes are affected by the previous nodes. Therefore, the forward calculation formula of the hyper-DAG is relatively complex. For convenience, the forward computation of each cell is simplified as:

$$f_{cell}(x) = \underbrace{\begin{bmatrix} \alpha_0^0 & \cdots & \alpha_M^0 \\ \vdots & \ddots & \vdots \\ \alpha_0^{\frac{N(N+1)}{2}-1} & \cdots & \alpha_M^{\frac{N(N+1)}{2}-1} \end{bmatrix}}_{A} \otimes \underbrace{\begin{bmatrix} OP_0(x) \\ \vdots \\ OP_M(x) \end{bmatrix}}_{Search\ space} \quad (2)$$

where N is the number of intermediate nodes. All the α vectors are spliced into a coefficient matrix, denoted as A, and each row of this matrix represents a mixed operators edge. The symbol $\otimes$ is defined as the interaction of the coefficient matrix and the search space.

The secondary step is to optimize the coefficient matrix and select the operator with the maximum coefficient value from each mixed edge. As shown in the Step 2 of Fig. 1, there are three operators between N1 and N3 nodes, and the $\alpha_2^4 = 0.6$ operator is finally retained. Then, the final sub-DAG is obtained by selecting the maximum coefficient from each mixed edge.

The searching process can be regarded as a bilevel optimization problem, as shown in the following equation:

$$A^* = \arg\min_A L_{valid}(w^*(A), A)$$
$$\text{s.t. } w^*(A) = \arg\min_w L_{train}(w, A) \quad (3)$$

where w^* is the weights of filters under fixed A, and A^* denotes the optimized coefficient matrix. As a result, the searching process is reduced to the optimization of A.

C. Speed Up Strategy

To reduce computing costs, DARTS uses two speed up strategies: proxy model and bilevel optimization approximation. For the former, DARTS searches for a repeatable cell structure as mentioned above, and this strategy is also used in manually designed networks, such as ResNet and Inception. For the latter, DARTS alternates optimization by coupling the different optimization objectives of training and searching, which greatly speeds up the searching process at the expense of partial performance.

Additionally, to accommodate large datasets, DARTS uses a dataset proxy (i.e. a small dataset) to search the optimized cell, and then stacks the deeper network to apply in the large scale dataset. It is worth noting that this strategy has implication for transfer learning based diagnosis, and Li and Peng [22] proposed a domain adaptation method based on DARTS.

III. EXPERIMENTAL STUDY

In this section, the aeroengine bevel gear dataset is collected and used to verify the performance of DARTS. The comparison experiments are also carried out on the four commonly used models, namely the standard DCNN, VGG11, ResNet18 and Inception. The standard DCNN has a series stack of 5 convolutional layers. The VGG11, ResNet18 and Inception are directly copied from a benchmark study [23]. All the experiments are running on the Ubuntu 18.04 GNU/Linux and GeForce RTX 2080Ti.

A. Data Description

As shown in Fig. 2(a), the bevel gear is an important part of the aeroengine transmission system, so it is necessary to monitor its state. In the testing bench, as shown in Fig. 2(b), we set up four groups of bevel gear experiments in different states, each of which was operated at five rotational speeds respectively. The vibration signals were acquired by acceleration sensor, and the detection points are shown in Fig. 2(c). The states and speed information are described in Table II.

(a)   (b)   (c)
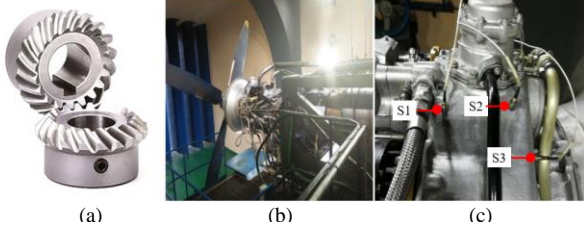
Fig. 2. The testing bench for aeroengine bevel gear. (a) Aeroengine bevel gear. (b) The testing bench. (c) Sensor position layout

TABLE II.   THE AEROENGINE BEVEL GEAR DATASET

Data Type	Health state	Broke Teeth	Tooth surface wear	Overhaul
Rotational speed	500/1000/1500/2000/3900rpm			
Samples for training	600*5	600*5	600*5	600*5
Samples for testing	300*5	300*5	300*5	300*5
Label	0	1	2	3

For each rotational speed, the ratio of train set and test set is 2:1, and one sample has 1024 points without data preprocessing or augmentation. Samples of different speeds in the same mode share the same label. Details are shown in Table II. Therefore, it can be seen as a multi-condition classification problem.

B. Noise Addition

To verify the performance of DARTS, different levels of Gaussian noise are added to the original signal according to the SNR equation described in Eq.(4), and the added SNR in this study are {-10dB, -5dB, 0dB, 5dB, 10dB} respectively.

$$SNR(dB) = 10\log_{10}(\frac{P_{signal}}{P_{noise}}) \qquad (4)$$

where P_{signal} is the signal power, P_{noise} is the noise power.

C. Results

During the searching process of DARTS, the cell is defined with two input nodes and four intermediate nodes. The search space, as shown in Table I, has seven operators and a special zero operator. Additionally, the same hyper-parameter strategy is applied to the training process of all the models, including SGD optimizer, cosine annealing learning rate and gradient clipping. In order to eliminate the randomness of the results, the final accuracy is the average of the last 10 epochs.

Multiple optimized sub-networks for different SNR datasets are searched based on DARTS. The sub-network corresponding to the -10dB SNR is shown in Fig. 3. It is worth noting that DARTS will search a customized network for each SNR dataset. According to the experiments, generally, only 3 GPU hours are required for DARTS to design high-performance network for this dataset.

The experimental results of DARTS and the commonly used models are shown in Table III. As can be observed from these results that with the increase of noise, the accuracy of all models decreases. The accuracy of the standard DCNN decreases the most, while that of DARTS decreases the least. However, we can see that DARTS outperforms other popular manually designed networks at the same noise level. Besides, DARTS is very competitive in model size compared with other networks, and the lightweight model has superiority when applied to the resource-constrained scenarios.

It is also interestingly found that ResNet18 and Inception perform less well than the standard DCNN under certain SNR. As mentioned above, the architectures of these two networks are tailored to datasets in the CV domain, which are suitable for processing images with rich spatial information. The particular tricks are used to fit the open source image datasets. For instance, ResNet and Inception both use specific parameter initialization methods that are statistically significant for image. More directly, ResNet will initialize some special layers with special parameters to improve accuracy by 0.2% on the open source image dataset. Obviously, these tricks are not suitable for the vibration signal, and the result shows that the performance of these two network architectures is also unsatisfactory. In fact, it is more appropriate to learn the core concepts of these network architectures, such as the identity mapping of ResNet and the multi-solution channel of Inception. Then, NAS provides an efficient way to design task-specific network from these concepts.
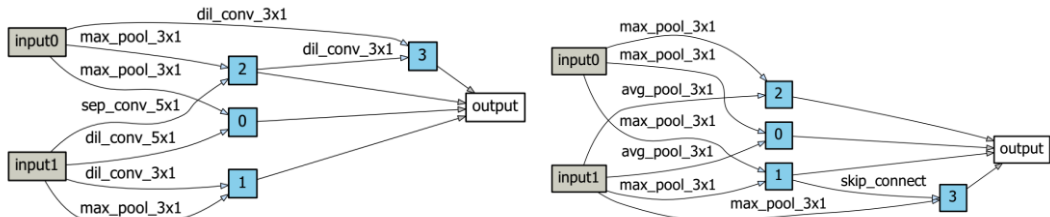
Fig. 3. The normal cell (left) and reduce cell (right) for the dataset with -10dB SNR. For each cell, {input0, input1} are the input nodes, {0, 1, 2, 3} are the intermediate nodes and {output} is the output node. It is worth noting that there are 2, 3, 4 and 5 edges for intermediate nodes respectively. After further filtering, each intermediate node retains only two selected edges with the maximum coefficient from the mixed edges.

TABLE III.   THE EXPERIMENTAL RESULTS

Model Name	Model Size	SNR Level					Noise free
		-10	*-5*	*0*	*5*	*10*	
DCNN	6.4M	36.20(0.15)	62.79(0.07)	87.02(0.01)	96.16(0.01)	99.09(0)	99.67(0)
VGG11	31.3M	44.49(0.09)	65.74(0.06)	88.02(0.02)	96.53(0.1)	99(0.01)	99.59(0.01)
ResNet18	15.4M	44.37(0.13)	65.28(0.04)	86.08(0.05)	94.99(0.02)	98.11(0.01)	99.37(0.01)
Inception	3.3M	48.35(0.12)	64.72(0.08)	82.21(0.03)	92.76(003)	96.58(0.02)	99.22(0)
DARTS	6.3M	**50.85**(0.65)	**69.69**(0.15)	**90.32**(0.27)	**97.19(0.04)**	**99.49**(0)	**99.71**(0.03)

IV. CONCLUSION

This paper proposes a differentiable NAS method for aeroengine bevel gear fault diagnosis, which is very efficient and resource-friendly to design a customized network for a specific dataset. This NAS method is applicable not only to the trunk model for fault diagnosis, but also to other sub-tasks of PHM. In addition, the experimental results show that our method is more effective than other manually designed models in fault diagnosis under low SNR.

ACKNOWLEDGEMENT

REFERENCES

[1] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring sased on Hilbert–Huang Transform, support vector machine, and regression," IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 1, pp. 52–62, 2015, DOI: 10.1109/tim.2014.2330494.

[2] C. Li, R. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals," Mechanical Systems and Signal Processing, vol. 76-77, pp. 283–293, 2016, DOI: 10.1016/j.ymssp.2016.02.007.

[3] T. Li, Z. Zhao, C. Sun, R. Yan and X. Chen, "Adaptive Channel Weighted CNN With Multisensor Fusion for Condition Monitoring of Helicopter Transmission System," IEEE Sensors Journal, vol. 20, no. 15, pp. 8364-8373, 2020.

[4] Z. Hu, Y. Wang, M. Ge, and J. Liu, "Data-driven fault diagnosis method based on compressed sensing," IEEE Transactions on Industrial Electronics, vol. 67, pp. 3216-3225, no. 4, 2020, DOI: 10.1109/tie.2019.2912763.

[5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," In International Conference on Learning Representations, 2015.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," In Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2015.

[7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V.Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[8] X. Chu, B.Zhang, R. Xu, and J. Li, "FairNAS: rethinking evaluation fairness of weight sharing neural architecture search," 2019, arXiv: 1907.01845.

[9] H. Xu, L. Yao, W. Zhang, X. and Liang, Z. Li, "Auto-FPN: automatic network architecture adaptation for object detection beyond classification," IEEE International Conference on Computer Vision, pp. 6648-6657, 2019.

[10] A. Piergiovanni, A. Angelova, A. Toshev, and M. S. Ryoo, "Evolving space-time neural architectures for videos," IEEE International Conference on Computer Vision, pp. 1793-1802, 2019.

[11] N. Klyuchnikov, I. Trofimov, E. Artemova, M. Salnikov, M. Fedorov, and E. Burnaev, "NAS-Bench-NLP: neural architecture search benchmark for natural language processing," 2020, arXiv: 2006.07116v1.

[12] A. Baruwa, M. Abisiga, I. Gbadegesin, and A. Fakunle, "Leveraging end-to-end speech recognition with neural architecture search," 2019, arXiv: 1912.05946v1.

[13] R. Wang, H. Jiang, X. Li, and S. Liu, "A reinforcement neural architecture search method for rolling bearing fault diagnosis," Measurement, vol. 154, 2020, DOI: 10.1016/j.measurement.2019.107417.

[14] X. Li, Y. Hu, J. Zheng, and M. Li, "Neural architecture search for fault diagnosis," 2020, arXiv: 2002.07997v1.

[15] J. Li and D. He, "A Bayesian optimization AdaBN-DCNN method with self-optimized structure and hyperparameters for domain adaptation remaining useful life prediction," IEEE Access, vol. 8, pp. 41482-41501, 2020, DOI: 10.1109/access.2020.2976595.

[16] B. Zoph and Q. V. Le, "Neural architecture search with reinforcement learning," In International Conference on Learning Representations, 2017.

[17] E. Real, A. Aggarwal, Y. Huang, and Q. V. Le, "Regularized evolution for image classifier architecture search," In AAAI conference on Artificial Intelligence, 2019.

[18] H. Pham, M. Y. Guan, B. Zoph, Q. V. Le, and J. Dean, "Efficient neural architecture search via parameter sharing," In International Conference on Machine Learning, 2018.

[19] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," In Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, 2018.

[20] R. Luo, F. Tian, T. Qin, E. Chen, and T. Liu, "Neural architecture optimization," In Conference on Neural Information Processing Systems, 2018.

[21] H. Liu, K. Simonyan, and Y. Yang, "DARTS: differentiable architecture search," In International Conference on Learning Representations, 2019.

[22] Y. Li and X. Peng, "Network architecture search for domain adaptation," 2020, arXiv: 2008.05706v1.

[23] Z. Zhao, T. Li, and J. Wu, et al. "Deep Learning Algorithms for Rotating Machinery Intelligent Diagnosis: An Open Source Benchmark Study," ISA Transactions 2020, DOI: 10.1016/j.isatra.2020.08.010

Rolling Bearing Fault Diagnosis Based on Horizontal Visibility Graph and Graph Neural Networks

Chenyang Li
School of Instrument Science and Engineering
Southeast University
Nanjing, China
licy@seu.edu.cn

Lingfei Mo*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
lfmo@seu.edu.cn

Ruqiang Yan*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
ruqiang@seu.edu.cn

Abstract—The automatic extraction and learning features relying on artificial intelligence algorithms replace traditional manual features. More effective feature expression improves the performance of machine fault diagnosis with fewer requirements for labor and expertise. However, the present models only can process the data in Euclidean space. The relations between data points are ignored for a long time, which can play a significant role in distinguishing diverse faults patterns. To combat this issue, a novel model for bearing faults diagnosis is proposed by incorporating the horizontal visibility graph (HVG) and graph neural networks (GNN). In the proposed model, time series is converted to graph retaining invariant dynamic characteristics through the HVG algorithm, and the generated graphs are fed into a designed GNN model for feature learning and faults classification further. Finally, the proposed model is tested on two actual bearing datasets, and it shows state-of-the-art performance in the bearing faults diagnosis. The experimental results demonstrate that extracting relation information using HVG benefits bearing faults diagnosis.

Keywords— rolling bearing, fault diagnosis, horizontal visibility graph (HVG), graph neural networks (GNN)

I. INTRODUCTION

With mechanical equipment becoming larger, high-speed, more precise, and high automation, the impact and harm caused by equipment failure is getting more and more serious, such as equipment damage, system collapse, and production stagnation. Therefore, the health management of mechanical equipment in intelligent manufacturing has become one of the key problems to be solved urgently in the system intelligence.

Traditional mechanical fault signal processing methods mainly focus on the time domain, frequency domain, and time-frequency domain, which can combine with machine learning methods such as multi-layer perceptron (MLP), support vector machine (SVM), and Bayesian estimation to diagnose the data with obvious features and simple patterns. With the rapid development of artificial intelligence, intelligent algorithms represented by deep learning have gradually emerged in the area of fault diagnosis [1], for instance, convolutional neural networks (CNN) [2], [3], automatic encoders (AE) [4], generative adversarial networks (GAN) [5], etc. Current feature extraction algorithms are aimed at the data value itself, but the ability to extract relations and structure between the data is limited.

In recent years, there is a growing interest in analyzing time series using complex networks in some specific domains. The common methods for transforming time series into graph representation are recurrence networks, transition networks, and visibility graphs [6] in terms of different principles. The application of the graph model in fault diagnosis has just getting started. Gao et al. divided the non-stationary periodic signal into a single period and utilized the graph structure to detect the change of mechanical equipment's working states [7]. Gao et al. transformed the original vibration signal to graph domain, and used graph Fourier transform to extract the graph spectrum coefficients for fault diagnosis [8]. Furthermore, the total variation which reflects the smoothness of graph signals served as the indicator to distinguish different rolling bearing faults signals [9].

Unlike the data in Euclidean space, the data in non-Euclidean space contain the complex structure and abundant relation information, such as the physical system, molecular fingerprint, social networks, recommendation systems, etc. The relational data pose a great challenge for previous machine learning and deep learning algorithms. Graph neural networks (GNN) emerging as a specialized graph signal processing algorithm have gained extensive attention from researchers [10]. GNN can achieve more effective data relationship extraction and inference by aggregating the information of the node's neighbors at any depth. At present, the main GNN models can be classified into four categories: recurrent GNN, convolutional GNN, graph auto-encoders, and spatial-temporal GNN [11]. GNN has been successfully applied in the field of physical models [12], chemical structure [13], social networks [14], natural language processing [15], image classification [16].

As a deep learning method based on graph domain, GNN has convincing performance and high interpretability in graph data learning, but there is barely any research about the application of GNN in intelligent diagnosis of mechanical equipment. In this paper, we devote to explore the feasibility of GNN in rolling bearing fault diagnosis. By the HVG algorithm, the original vibration data are converted to a whole graph composed of the numerical value and connections between the data. A GNN model is designed for the graph classification referring to the state-of-the-art model GIN. Experiments on two bearing datasets are conducted to verify the effectiveness of GNN in fault diagnosis. Besides, the comparison with MLP, recurrent neural networks (RNN), and

(a) Time series histogram

(b) HVG
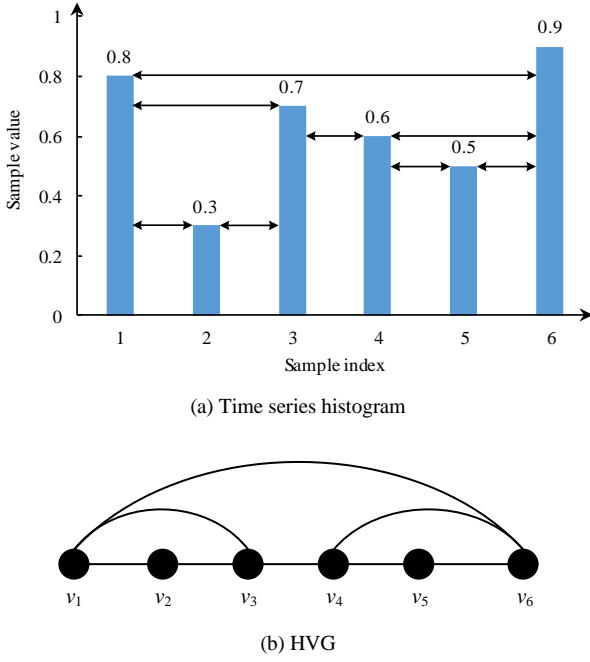
Fig. 1. Illustration of HVG algorithm

its variants are given to demonstrate the improvement causing by the introduction of data relations.

II. THEORETICAL BACKGROUND

The HVG algorithm proposed by Luque et al. is based on the visibility graph [17], [18]. The HVG algorithm maps the time series into graph in non-Euclidean space based on the horizontal relationship between data points. A graph G can be denoted as $G = (V, E)$, where $V=\{v_1, v_2, …, v_N\}$ is the set of N nodes, $E=\{e_{ij}| v_i$ and v_j are linked$\}$ is the edge set. Suppose a time series has N data points with values $X=\{x_1, x_2, …, x_N\}$, and the nodes are the discrete data points with the feature of corresponding data value. According to the horizontal visibility criterion, two nodes are considered linked only if a horizontal line connects x_i and x_j without intersecting by any intermediate data height, i.e., the following geometrical relationship is satisfied

$$x_i, x_j > x_n, \ \ \text{all } n \in (i, j). \tag{1}$$

An illustration of the algorithm is given in Fig. 1. The top shows the histogram of a sampled time series {0.8, 0.3, 0.7, 0.6, 0.5, 0.9}. The edge exists when the bar heights of the two nodes are higher than any other data heights between them. By connecting the bars and the horizontal lines as the nodes and edges, the HVG can be generated as shown in Fig. 1(b). Naturally, the HVG possesses connectivity, invariance under affine transformations, and other characteristics. Graph neural networks is a feasible approach to deal with the HVG.

GNN, as a rising neural network inspired by Graph Embedding and CNN, shows a powerful ability to learn graph in the non-Euclidean space. Passing messages along edges can obtain representation and extract features of the nodes and graph [19]. A graph $G = (V, E)$ with N nodes has a node feature matrix $X \in \mathbb{R}^{N \times F}$, where F is the dimension of the node feature vector. GNNs embed the node feature representing as a hidden state h_v along the graph topology. The forward pass of current GNNs models follows the aggregation and readout strategy. In the first phase, the node features are updated by aggregating the information of its neighbors, the k-th propagation is formulated as [19], [20]:

$$m_v^{(k)} = \text{AGGREGATE}^{(k)}\left(\left\{h_u^{(k-1)} : u \in \mathcal{N}(v)\right\}\right) \tag{2}$$

$$h_v^{(k)} = \text{COMBINE}^{(k)}\left(h_v^{(k-1)}, m_v^{(k)}\right) \tag{3}$$

where $m_v^{(k)}$ is the passing message generated by adjacent nodes, $\mathcal{N}(v)$ is the neighbor set of node v, $h_v^{(k)}$ is the output feature of node v in the k-th layer, and $h_v^{(0)}$ is initialized with x_v. Through the k iterations of aggregation, the node feature can fuse the information within its k-hop neighbors, i.e., the receptive filed reaches to k-hop neighbors. For node-level tasks, the node label y_v can be inferred by the final node representation: $y_v = f(h_v)$. For the graph-level tasks, a graph representation h_G is further required, which will be computed in the second phase through a READOUT function:

$$h_G = \text{READOUT}\left(\left\{h_v^{(k)} \mid v \in G\right\}\right). \tag{4}$$

The choices of the AGGREGATE function, COMBINE function, and READOUT function have a great effect on the learning ability of GNNs. Apart from differentiable characteristic, the READOUT function also needs to be invariant to node permutations. Most representative GNNs models can be reformulated to the aggregation framework. The propagation rule of Graph Convolutional Networks (GCN) [21] utilizes the first-order approximation of spectral graph convolution:

$$h_v^{(k)} = \text{ReLU}\left(W \cdot \text{MEAN}\left\{h_u^{(k-1)}, \forall u \in N(v) \cup \{v\}\right\}\right) \tag{5}$$

where W is a learnable matrix, MEAN denotes the mean aggregator. GraphSAGE, an inductive framework, analyzed the Mean aggregator, LSTM aggregator, and Pooling aggregator [14]. The Pooling aggregator had the best performance in experiments, where a max-pooling operator served as the aggregator:

$$m_v^{(k)} = \text{MAX}\left(\left\{\text{ReLU}\left(W \cdot h_u^{(k-1)}\right), \forall u \in \mathcal{N}(v)\right\}\right). \tag{6}$$

Motivated by the Weisfeiler-Lehman (WL) graph isomorphism test, Xu et al. point out that the upper bound of the GNNs based on nodes aggregation is the WL test if the aggregator and graph-level readout functions are injective [20]. Based on this theorem, a simple but effective architecture, Graph Isomorphism Network (GIN) is developed to capture dependencies of isomorphic graphs. Considering the universal approximation of MLP, MLP is used to model the aggregation and combine functions:

$$h_v^{(k)} = \text{MLP}^{(k)}\left(\left(1+\varepsilon^{(k)}\right) \cdot h_v^{(k-1)} + \sum_{u \in N(v)} h_u^{(k-1)}\right) \tag{7}$$

where ε is a learnable or fixed parameter. Because the sum operator is injective, it is selected for aggregator instead of max or mean. GIN achieves state-of-the-art performance on several graph classification benchmarks.

III. MODEL ARCHITECTURE

Normally, the working condition is monitored by sensors, thus, the potential fault information lies in the sampled time series. Through the HVG algorithm, we can transform a time series sample into a graph with condition-specific topology,
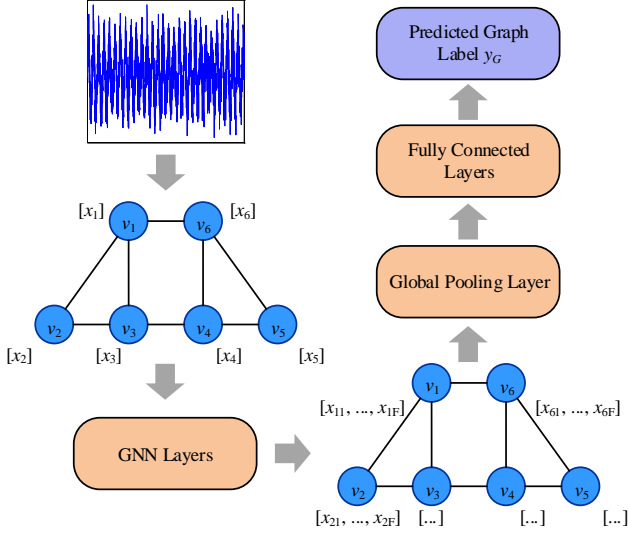
Fig. 2. The proposed GNNs architecture for fault diagnosis

and the sampling point value is treated as the corresponding node's feature. In this way, the fault diagnosis is converted to a graph-level classification task. To obtain the graph feature, we design a deep GNNs architecture based on the GIN model, which consists of three main parts: GNN layers, a global pooling layer, and two fully connected (FC) layers. The illustration of the GNNs architecture is shown in Fig. 2.

The GIN aggregator has been adjusted to better adapt to bearing faults data, but still employs injective functions. The MLP is replaced by 1-layer perceptron, i.e., a linear mapping function. While the representation ability of the 1-layer perceptrons is considered insufficient, it brings the equivalent performance as well as a stable training process on our bearing datasets. The sum aggregator is chosen the same as GIN and ε is fixed to 0. Thus, the aggregation of the GNN layer is defined as follows:

$$h_v^{(k)} = \text{Linear}^{(k)}\left(h_v^{(k-1)} + \sum_{u \in N(v)} h_u^{(k-1)} \right). \quad (8)$$

After passing the nonlinear activation function ReLU and a batch normalization layer, the aggregated graph is input into the next aggregation layer. By stacking multiple aggregation layers, the node feature can merge with the graph structure. Then, to get the entire graph representation r_G, the node features go to a global sum pooling layer:

$$r_G = \sum_{i=1}^{N} h_{v_i}. \quad (9)$$

After adding node features across the node dimension, one can get the entire graph feature to infer the fault type. The graph label is finally inferred by two FC layers with softmax function.

The negative log-likelihood (NLL) loss function calculates the negative log value of probability predicted by the softmax layer:

$$L(y_G) = -\log(y_G) \quad (10)$$

where y_G is the predicted label.
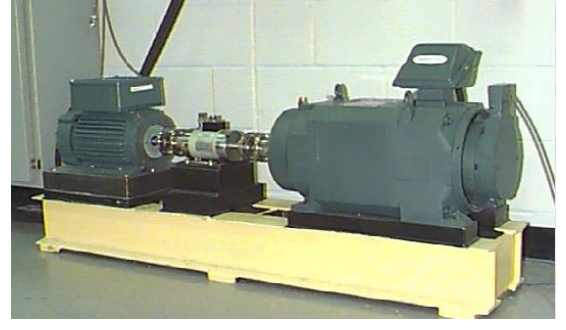
Fig. 3. CWRU bearing test rig [22]

TABLE I.   CRWU BEARING FAULT TYPES DESCRIPTION

Types	Description	Sample size
Normal	Healthy bearing	100
BE	Ball element faults with 3 fault diameters	100*3
IR	Inner race faults with 3 fault diameters	100*3
OR	Outer race faults with 3 fault diameters	100*3

IV. EXPERIMENTS

To validate the effectiveness of the proposed method, we perform the experiments on two actual bearing datasets: Case Western Reserve University (CWRU) bearing dataset and Southeast University (SEU) bearing dataset. The comparative experiment with MLP is performed to highlight the influence of the latent structure of the time series. Moreover, we also analyze the feature extraction ability of different numbers of aggregation layers.

A. Datasets

1) CWRU bearing dataset

CWRU bearing dataset is a public dataset for bearing faults diagnosis, which has been tested a lot in previous research [22]. The bearing experiments were carried out on a 2 hp electric motor, and the vibration data were collected using accelerometers deployed on both the drive end and fan end of the motor housing as shown in Fig. 3. The faults were set manually by electro-discharge machining with different damage intensity. In our experiments, data in load 2 hp with three fault diameters: 0.1778 mm (0.007 inches), 0.3556 mm (0.014 inches), 0.5334 mm (0.021inches) are selected, and each fault diameter corresponds to three fault types. Thus, it becomes a ten-class classification task, and each class contains 100 samples. 1024 data points are partitioned to one sample with the sampling rate of 12000 Hz. The data adopted in the following experiments are listed in TABLE I.

2) SEU bearing dataset

The SEU bearing fault simulation data are collected on Drivetrain Dynamics Simulator (DDS). As depicted in Fig. 4, DDS consists of a motor, a planetary gearbox, a parallel gearbox, a brake, and controller modules, which can simulate different kinds of faults of gearboxes and bearings. The acceleration sensors are mounted on the drive motor, planetary gearbox, and parallel gearbox, respectively. Setting the rotating speed and load to 30 Hz and 2 V, the data of each bearing fault type are acquired at a sampling rate of 5120 Hz.
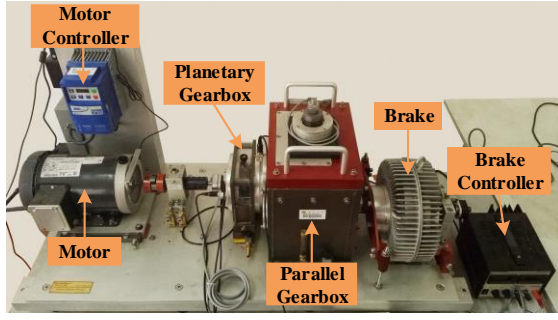
Fig. 4. SEU gearbox test rig

TABLE II.   SEU BEARING FAULT TYPES DESCRIPTION

Types	Description	Sample size
Normal	Healthy bearing	800
Ball	Ball element fault	800
Inner	Inner ring fault	800
Outer	Outer ring fault	800
Combination	Inner and outer ring faults	800

The same as CRWU datasets, a sample contains 1024 sampling points, and four types of failures and one health condition constitute the five-class classification task. The detailed information of the fault types is listed in TABLE II.

B. Experiment Configuration

In the experiments, at most 4 aggregation layers are stacked, followed by one global sum pooling layer and 2 FC layers for feature mapping and probability calculation. To analyze the improvement brought by the graph structure, we use MLP with equal layers as the baseline. The layer type and numbers are used to identify the neural networks architecture, such as 3GNN-2FC denoting 3 GNN layers and 2 FC layers.

Batch normalization is incorporated after every hidden layer. The hyperparameters are tuned through repeated experiments. The hidden unit size in the GNN model is determined to be 16, while MLP is 128. Adam optimizer with learning rate 0.005 and batch size 32 remain unchanged in both two datasets. The data are split for training, validation, and test with the percentages of 0.8, 0.1, 0.1. The training process stops when the validation accuracy does not improve for 10 epochs. For an accurate comparison, the final results average the accuracies of 10 repeated experiments with random data splitting. The model is implemented using Pytorch Geometric libraries [23].

C. Results

Through transforming time series to graph using the HVG algorithm, the topology relation among the data can be introduced into the data. To investigated whether the HVG has extracted extra useful information, the proposed GNN designed for learning graph data is compared with the MLP model where only the samples' numerical values are input. The classification accuracies of the designed model and MLP are shown in Fig. 5 and Fig. 6 (CRWU and SEU bearing dataset). Intuitively, the proposed GNN model outperforms the MLP with equal layers. Stacking more MLP layers, there is no obvious increase in classification accuracy. However, the

aggregator can capture longer dependency of nodes and obtain more helpful features using deeper GNN layers. The GNN model has a satisfying performance even only containing a single aggregation layer. The fault diagnosis ability is improved greatly because of the introduction of geometry between data.

RNN has been studied extensively as an effective solution to time series analysis [24]. Several variants are proposed based on RNN, such as gated recurrent units (GRU) networks, bidirectional gated recurrent units (BiGRU) networks, and local feature-based gated recurrent units (LFGRU) networks [25]. These models can capture the temporal dependency
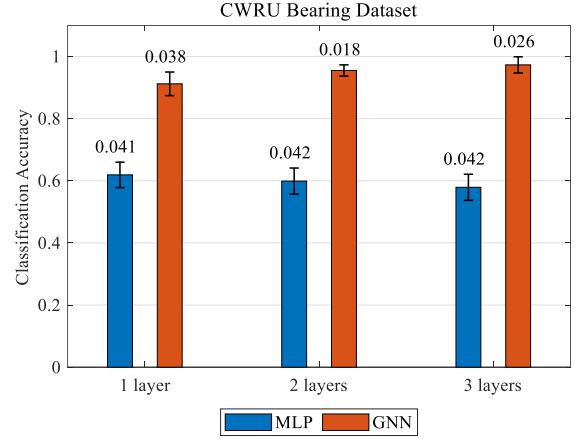
Fig. 5. Average accuracy and standard deviation of 10 repeated experiments on CWRU bearing dataset
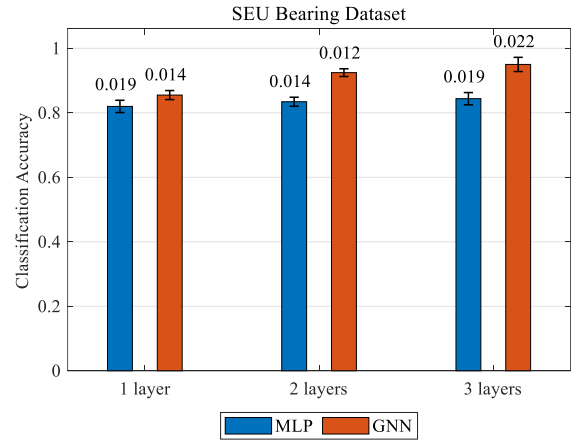
Fig. 6. Average accuracy and standard deviation of 10 repeated experiments on SEU bearing dataset

TABLE III.   CLASSIFICATION ACCURACY COMPARISON ON SEU BEARING DATASET

Models	Accuracy
RNN [25]	0.920
GRU [25]	0.924
BiGRU [25]	0.936
LFGRU [25]	0.940
3GNN-2FC (Ours)[a]	0.950
4GNN-2FC (Ours)	0.974

[a.] 3GNN-2FC denotes 3 GNN layers and 2 FC layers

between input data. Further, we compare the presented model with RNN and its variants, as exhibited in TABLE III. The proposed model performs better than RNN baselines when stacking 3 GNN layers and 4 GNN layers. The improvement attributes to the strong data representation ability of HVG and GNN.

V. CONCLUSION

A new model for bearing faults diagnosis based on HVG and GNN is presented in this paper. The original data series are first transformed to graphs in non-Euclidean space according to the visibility between data points. To extract the underlying features of the graph, GIN is modified to learn the graph feature generated from the vibration signal. Two bearing datasets are involved to validate the performance of the proposed method. The experimental results show that the transformation of HVG provides extra useful information for the classification comparing to pure numerical information. The GNN model also has superiority over RNN, and its variants. In future works, the weighted HVG and limited penetrable HVG will be investigated, and a more efficient GNN architecture will be explored to fit the characteristics of HVG.

REFERENCES

[1] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," Mech. Syst. Signal Process., vol. 115, pp. 213–237, 2019.

[2] L. Wen, X. Li, L. Gao, and Y. Zhang, "A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method," IEEE Trans. Ind. Electron., vol. 65, no. 7, pp. 5990–5998, 2018.

[3] Z. Zhu, G. Peng, Y. Chen, and H. Gao, "A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis," Neurocomputing, vol. 323, pp. 62–75, 2019.

[4] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification," Measurement, vol. 89, pp. 171–178, 2016.

[5] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," Comput. Ind., vol. 106, pp. 85–93, 2019.

[6] Z. Yong, D. R. V., M. Norbert, K. Jürgen, and D. J. F, "Nonlinear time series analysis by means of complex networks," Sci. Sin. Phys. Mech. Astron., vol. 50, no. 1, p. 010509, Jan. 2020.

[7] Z. Gao, G. Lu, and P. Yan, "Graph-based change detection for condition monitoring of industrial machinery: an enhanced framework for non-stationary condition signals," Meas. Sci. Technol., vol. 30, p. 115002, 2019.

[8] Y. Gao, D. Yu, and H. Wang, "Fault diagnosis of rolling bearings using weighted horizontal visibility graph and graph Fourier transform," Measurement, vol. 149, pp. 1–12, 2020.

[9] Y. Gao and D. Yu, "Total variation on horizontal visibility graph and its application to rolling bearing fault diagnosis," Mech. Mach. Theory, vol. 147, p. 103768, 2020.

[10] J. Zhou et al., "Graph Neural Networks: A Review of Methods and Applications," pp. 1–20, Dec. 2018, ArXive ID: 1812.08434.

[11] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," IEEE Trans. Neural Networks Learn. Syst., pp. 1–21, 2020.

[12] P. Battaglia, R. Pascanu, M. Lai, D. Jimenez Rezende, and K. Kavukcuoglu, "Interaction Networks for Learning about Objects, Relations and Physics," in Advances in Neural Information Processing Systems (NIPS 2016), 2016, pp. 4502–4510.

[13] A. Fout, J. Byrd, B. Shariat, and A. Ben-Hur, "Protein interface prediction using graph convolutional networks," in Advances in Neural Information Processing Systems Conference (NIPS 2017), 2017, pp. 6530–6539.

[14] W. L. Hamilton, Z. Ying, and J. Leskovec, "Inductive Representation Learning on Large Graphs," in Advances in Neural Information Processing Systems (NIPS 2017), 2017, pp. 1024–1034.

[15] Y. Zhang, Q. Liu, and L. Song, "Sentence-state LSTM for text representation," in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Long Papers), 2018, vol. 1, pp. 317–327.

[16] X. Wang, Y. Ye, and A. Gupta, "Zero-Shot Recognition via Semantic Embeddings and Knowledge Graphs," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, pp. 6857–6866.

[17] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, "From time series to complex networks: The visibility graph," in Proceedings of the National Academy of Sciences of the United States of America, 2008, vol. 105, no. 13, pp. 4972–4975.

[18] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal visibility graphs: Exact results for random time series," Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys., vol. 80, p. 046103, 2009.

[19] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural Message Passing for Quantum Chemistry," in Proceedings of the 34th International Conference on Machine Learning, 2017, pp. 1263–1272.

[20] K. Xu, S. Jegelka, W. Hu, and J. Leskovec, "How powerful are graph neural networks?," in 7th International Conference on Learning Representations (ICLR 2019), 2019, pp. 1–17.

[21] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," in 5th International Conference on Learning Representations (ICLR 2017), 2017, pp. 1–14.

[22] "Case western reserve university bearing data center." [Online]. Available: http://csegroups.case.edu/bearingdatacenter.

[23] M. Fey and J. E. Lenssen, "Fast Graph Representation Learning with PyTorch Geometric," in 7th International Conference on Learning Representations (ICLR2019), 2019, pp. 1–9.

[24] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent Advances in Recurrent Neural Networks," 2017, ArXive ID: 1801.01078v3.

[25] R. Zhao, D. Wang, R. Yan, K. Mao, F. Shen, and J. Wang, "Machine health monitoring using local feature-based gated recurrent unit networks," IEEE Trans. Ind. Electron., vol. 65, no. 2, pp. 1539–1548, 2018.

Decoupled Feature-Temporal CNN: Explaining Deep Learning-Based Machine Health Monitoring

Rui Zhao
Quantitative Research
Harveston Asset Management
Singapore, Singapore
rui91seu@gmail.com

Zhenghua Chen*
Institute for Infocomm Research
A*STAR
Singapore, Singapore
chen0832@e.ntu.edu.sg

Chaoyi Zhu
School of Instrumentation and
Optoelectronic Engineering
Beihang University
Beijing, China
z00419@buaa.edu.cn

Zery Chan
Institute for Infocomm Research
A*STAR
Singapore, Singapore
zerychan1234@gmail.com

Jinjiang Wang
School of Mechanical and
Transportation Engineering
China University of Petroleum
Beijing, China
jwang@cup.edu.cn

Ruqiang Yan
School of Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
yanruqiang@xjtu.edu.cn

Abstract—**Machine learning, especially deep learning, have been extensively applied and studied in the area of machine health monitoring in the recent "big data" era. For machine health monitoring systems (MHMS), the major efforts have been put in designing and deploying more and more complex machine learning models which are also called as black-box models considering they are non-transparent towards their working mechanism. However, this research trend brings huge potential risk in real life. Since machine health monitoring itself is a high-stake decision scenario, the decision of the autonomous monitoring system should be trustworthy and reliable, which refers to obtain explainability. Then, it comes to the following key question: Why the deployed MHMS predict what they predict. In this paper, we shed some lights on this meaningful research direction: explainable machine health monitoring systems (EMHMS). In EMHMS, the machine doctor could act like a real doctor who can only make diagnosis but also describe the patient's symptoms. First, we propose a specific convolutional neural network (CNN) structure, named as DecouplEd Feature-Temporal CNN (DEFT- CNN), balances precision-explainability trade-off. Specifically, features and temporal information will be encoded in different stages of our model. Second, to explain the decision of the model, we adopt the gradient-based methods to generate features and temporal saliency maps highlighting which kind of features and time steps are key for the model's predictions. At last, we conduct the experimental studies in the CWRU bearing fault diagnosis dataset to verify the effectiveness of our proposed framework.**

Keywords—*machine health monitoring, deep learning, convolutional neural network, explainable machine learning*

I. INTRODUCTION

In the field of machine health monitoring, deep learning systems built upon industrial big data have been pervasive across various applications [1]–[4]. For example, Georg et al. adopted deep neural network for condition monitoring in wind turbine [5]. Xiao et al. combined recurrence quantification analysis and long short-term memory (LSTM) for fault diagnosis of induction motors [6]. To classify online tool wear during dry machining, Terrazas et al. utilized convolutional neural network (CNN) based real-time cutting force measurements [7]. In addition to diagnosis, Chen et al. proposed a novel attention-based deep learning model for machine remaining life prediction [8]. Li et al. [9] proposed a CNN with time-frequency domain features for rolling bearing's RUL prediction. All these previous studies validate the effectiveness of deep learning systems in the application of machine health monitoring. Although the reported performance metrics such as training and testing accuracies in diagnosis tasks are very high and convincing, the deployment of these powerful deep learning algorithms in real industrial environments have been lagging compared to other applications such as recommendation in E-commerce platforms and machine translations. It is because that machine health monitoring is one of high-stake decision scenarios where any potential mistakes may lead to catastrophic outcomes. Since these complex models cannot provide any explanations for users to understand models' predictions (e.g., why the model predicts a inner race fault for a gearbox), users cannot rely on and trust these machine health monitoring systems. In contrast to these deep learning based MHMS, traditional machine health monitoring and fault diagnosis techniques can provide intuitive explanations and justify the models' predictions [10], [11]. For example, signal processing techniques can be adopted to extract some metrics and the abnormal working condition may be inferred based on the value range of pre-defined working circumferences. However, these previous approaches have limited precision and scalability due to the requirement of intensive expert efforts.

In this paper, we propose a novel approach named DecouplEd Feature-Temporal CNN (DEFT-CNN) to solve the above trade-off problem between precision and explainability in the application of machine health monitoring. The proposed framework is applicable for all machine health monitoring tasks where the input data are vibration sensory data. For each data sample, local features extraction is conducted firstly to obtain a sequence of feature vectors which contain two directions: features and temporal ones, as shown in Figure1. The details will be presented in the following sections. In computer vision, the conventional CNN has shown powerful capability to learn meaningful representations from images by capturing spatial relationships of pixels across rows and columns of an image. In our proposed DEFT-CNN structure, there are two kinds of convolutional layers: one is to focus on feature direction while the other one is to focus on time axis, which is also called 1D convolutional layer. The decoupled domain-temporal modelling design is able to provide separate explanation focusing on features and temporal information. In our framework, one of the gradient-based methods named Gradient-weighted Class Activation Mapping (Grad-CAM) is adopted to explain the model decisions [12]. In our proposed CNN structure with Grad-CAM, features and temporal saliency maps can be visualized and pointed out which

---

features and time segments are discriminative over model predictions. What is more, users are able to gain insights into the inner information flow in the deep learning model and find directions for model optimization, bias removal and model debugging. At last, the trust of the users could be built by providing explanations of the model decisions, which is vital for the application of deep learning systems in the field of machine health monitoring. The contributions of our work could be summarized as the following aspects:

- Introduce one gradient-based deep learning explanation method, i.e., Grad-CAM, to understand the contributions from features and temporal directions in machine health monitoring tasks.
- Design the specific network structure of CNN, i.e., DEFT-CNN, so that the attributions from domain features and time dynamics could be analyzed and understood separately.
- The proposed framework achieves a good balance between precision and explainability and is applicable to all machine health monitoring tasks that the input data are sensory time-series data.

II. RELATED WORKS

In this section, a brief review of deep learning explanation techniques and the relevant applications on machine health monitoring systems is presented.

A. Explainable Machine Learning

The raid growth applications of machine learning, especially deep learning models in high-stake decision areas such as healthcare, finance and machinery health monitoring, raises the requirements of understanding complex models. For deep learning, there are two main categories of interpretable deep learning techniques: perturbation-based methods and gradient- based methods [13], [14]. Our adopted technique Grad-CAM belongs to the latter one: gradient-based methods.

Perturbation-based Methods: To capture the attributions of input features towards model predictions, the input data are perturbed and fed into the model to monitor how the model's predictions change. Zintgraf et al. modified the input image by moving a grey patch occluding part and visualized the locations of patches that are responsible for the correct predictions [15]. Most of these methods are not model-specific techniques so that the same technique can be applied to any kinds of machine learning models (not limited to deep learning models). Local interpretable model-agnostic explanations (LIME) proposed by Ribeiro et al. was aimed at explaining the predictions of any black-box classifiers based on the assumption that complex black-box models may be linear and simple locally [16]. The first step of LIME is to generate a new dataset containing permuted samples near the neighbor of the original samples and the corresponding predictions of the interested black-box models. Then, an intrinsic interpretable model such as decision tree or lasso regression model is trained over this new dataset. From the intrinsic interpretable model, the feature importance of the interested sample could be approximated. Except LIME, Scoped Rules (Anchors) and SHapley Additive exPlanations (SHAP) are other popular model-agnostic methods [17], [18].

Although perturbation-based methods have high flexibility that are applicable to all black-box models, their efficiency is limited especially when the input data are in a high dimensionality space.

Gradient-based Methods: Gradient-based methods are focusing on computed gradients during backpropagation to understand the information flow through neural networks. The gradients of errors with regards to each layer are used to update the weights of corresponding layer to make the predicted value closer to the ground truth. What is more, the gradients of the error regrading to the input features are related to the contributions for all input features: the larger the gradient, the more informative of the corresponding input features toward the specific output. The intuitive explanation lies that those gradient values describe how each input feature should change in order to make the model prediction closer to the label. As a result, the core idea behind gradient-based methods is that the magnitude of the gradients received by each input features with regard to a data point and its label can be used as a proxy to measure the importance of the underlying feature towards the identification of the label under consideration. However, these gradient values cannot be directly used as understandable measures of the input features. Different approaches corresponding to different post-processing and transformations of these gradients are proposed to turn them into meaningful attribution maps of the input features, which is also called salience maps [12], [19]–[21]. In our framework, Grad-CAM is selected considering that it is designed for CNN architectures and able to provide explanation in a very efficient way. In addition, our proposed framework is related to one recent work that also combined CNN and Grad-CAM to explain time series prediction [22]. In the application of machine health monitoring, the DEFT-CNN is built upon local feature extraction from time, frequency and time-frequency domains specialized for vibration signals.

B. Explainable Machine Health Monitoring System

During the recent years, various state-of-the-art deep learning models have been applied in machine health monitoring. Among these previous works, some research shares simi- lar motivations with our works, i.e., exploring explainable machine health monitoring systems. Kraus et al. proposed a structured-effect neural network for remaining useful life (RUL) prediction, where model parameters are estimated by variational Bayesian inferences [23]. Compared to deep learning, Bayesian neural network is able to have a posterior inference that can capture uncertainty. Except Bayesian ma- chine learning that is good at uncertainty modelling, it cannot provide any explanations behind why the model makes such decisions. In [24], Lee et al. proposed a framework consisting of an autoencoder and a feedforward neural network to estimate the remaining useful life (RUL) of rotating machinery and the octave-band filtering was adopted to compress the sizes of the model and features to improve the explainability level. However, in their proposed framework, the features are limited to be power spectra of vibration signals due to the utilization of the octave-band filtering. John et al. adopted Layer-wise Relevance Propagation to explain the behavior of CNN that takes time-frequency spectra images as

inputs [25]. Different from their study, our proposed work is more flexible that can utilize time domain, frequency domain and time-frequency domain features.

III. Decoupled Feature-Temporal Cnn

In this section, our proposed DEFT-CNN is firstly presented. The key characteristic of the proposed structure is that the feature and temporal information are captured by different convolutional layers. It leads to the independent application of Grad-CAM on these two layers to estimate the contributions from feature and time dynamics instead of a joint effect.

Convolutional neural networks (CNNs) have been one of the most widely used deep learning models in the application of machine health monitoring since they are able to find patterns hidden in the input data. In the field of computer vision, there were previous works focusing on saliency maps that highlighting the important region in the input image for model prediction. However, it is challenging to utilize those frameworks directly in machine health monitoring, considering the input sensory data are time series data and most of them have multiple variate at each time step. Even though the input data are usually formed in a matrix, the two axis: feature and temporal directions should be considered independently. It motivates the development of the DEFT-CNN.

Similar to the work in [26], local feature extraction is applied on the sensory input. As shown in Figure 1, the original univariate time series data is transformed into a short sequence of feature vectors that usually consists of features in time, frequency and time-frequency domains. Then, the input data into the following CNNs is in the shape of $\mathbf{x} \in \mathbb{R}^{T \times d}$. Here, T and d denote the length of the sequence and the number of features at each time step.

As shown in the Figure 2, the DEFT-CNN learns abstract representations from features and temporal directions, respectively. The first convolution layer utilized a 2D filter whose filter size is. By setting the second dimension of the filter size to 1, the convolution operation is able to capture each feature independently. It enables the following gradient based method to compute the attention value of each single feature and generate the feature saliency map. The first 2D convolution layer is followed by a 1×1 convolutional layer which was firstly proposed in the inception architecture [27]. By setting the number of filter to 1, the dimensionality of feature map space could be reduced to 1 while the spatial dimension space is kept unchanged. The benefits are two-folds: first, the 1D convolutional layer could be applied on the output of the 1×1 convolutional layer; second, the reduced model complexity could overcome the potential overfitting problem. The third convolutional layer is a 1D convolutional layer, sliding the filter from the beginning of the time step to the ending of the time step. The filter will take all features in certain time steps and only differentiate time steps. Therefore, by computing the gradient flow of this convolutional layer, the temporal saliency map can be generated which could be used to explain which time segments are important for model prediction. After flatten, all generated feature maps are concatenated into a vector which can be fed into fully connected layers and a softmax layer as the top layers to predict targets.

The adopted gradient-based explainability technique here is named gradient-weighted class activation mapping (Grad-CAM) [12]. Grad-CAM utilizes the gradient flow from output neurons backpropagated into the convolutional layer to understand model predictions. Firstly, the gradients of the score for the predicted class y_c with respect to all features maps generated by a certain convolutional layer are computed. Here, c denotes the index of the predicted class while A^k represents the k-th feature map generated by the corresponding convolutional layer. Via global average-pooling operation over the spatial dimension of the feature map, the importance score of the feature map for the target class could be computed as follows:

$$a_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y_c}{\partial A_{ij}^k} \tag{1}$$

Then, with the computed scores as weighting terms, a weighted linear combination of feature maps could be computed and followed by the ReLU function as:

$$L^c = ReLU\left(\sum_k a_k^c A^k\right) \tag{2}$$

The ReLU function is able to filter features that have negative effects on the class of interest and render better visualization. By checking the output matrix L^c, the values in the areas of the input data are positively related to contributions toward the class of interest. Due to the decoupled structure in our proposed framework as shown in Figure 2, the Grad-CAM technique could be applied over the outputs of the first convolutional layer and the third 1D convolutional layer regarding to the class of interest, respectively. Then, feature saliency map and temporal saliency map could be generated to help us understand which features and time segments are most useful for model predictions. Therefore, our proposed framework is able to achieve a good balance-off between precision and explainability.
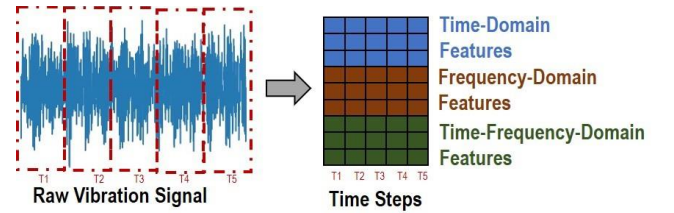
Fig. 1. Illustration of adopted local feature extractions.
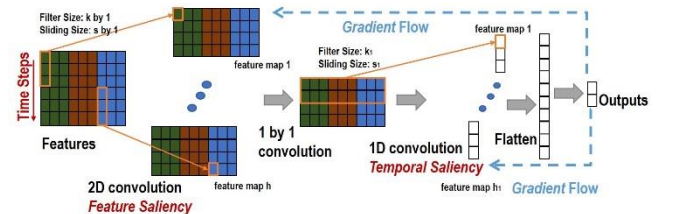
Fig. 2. Illustrations of the proposed DEFT-CNN. As shown in the figure, the numbers k, s and h are hyper-parameters for the first 2D convolutional layer. And the number k_1, s_1 and h_1 are hyper-parameters for the third 1D convolutional layer.

IV. EXPERIMENTS

In this section, our proposed framework is validated in one real dataset for rolling element bearing diagnosis from the Case Western Reserve University (CWRU) Bearing Data Center. The CWRU data has been a benchmark and standard dataset for machinery fault diagnostic algorithms. Due to the page limits, the detailed information of the CWRU dataset is skipped and one may find more information behind the dataset in [28]. Our implementation[1] has been public.

A. Implementation Details

As seen in Table I, three kinds of measures from each feature domain including time, frequency and time-frequency are selected so that nine features are designed. The length of the original input time series is 5012 and is then, sliced into 20 windows. Each sub-window of time series data whose length is 256 is transformed into a vector of nine measures. Therefore, each data sample is transformed into a matrix whose shape is 20 by 9. Then, this matrix will be used as the input into the following CNN model to predict the fault type of a bearing. A four-class dataset is prepared that four kinds of output labels are normal condition, inner race defect, outer race defect and ball defect. Each class contains 3592 data points and the total data size is 14368. After random splitting, 80% dataset is used as training data while the rest 20% dataset is regarded as testing data.

TABLE I.   LIST OF EXTRACTED FEATURES

Domain	Features (symbols)
Time	rms (rms) Peak-to-Peak (p2p) Skewness (skew)
Frequency	Spectral Kurtosis (*kurt) Spectral Skewness(*skw) Spectral power (*pow)
Time-Frequency	db1 Wavelet Energy (db1) db2 Wavelet Energy (db2) db3 Wavelet Energy (db3)

Since the data size in our experimental study is not very large, our proposed framework follows a "shallow" structure, as shown in Figure 2. In the first convolutional layer, the filter size and stride size are set to be $l_1 = (5, 1)$ and $s_1 = (1, 1)$. By setting the number in filter size and stride size along the feature axis to 1, the activation neuron in the generated feature map can be connected to each single feature. This unique connections enable the computed gradients could be used to explain the model decision influenced by each single feature. This convolutional layer contains 32 feature maps and adopts valid padding. Then, we adopt the Network in Network technique and fed the output of the first convolutional layer focusing on feature axis into the one by one convolutional layer with single filter. The output of the one by one convolutional layer will be in the shape of (16, 9). Then, the third convolutional layer will be a one-dimensional convolutional layer whose kernel size is $l_2 = 3$, stride size is $s_2 = 1$, number of feature maps is 32 and the mode of padding is valid. This layer is focused on modelling the temporal information. After flatten operation, the output of the last convolutional layer is fed into a dense layer whose hidden size is 64. For final classification, the last layer is set to be a dense layer with softmax activation and four output nodes corresponding to the number of classes. And the non-linear activation functions of all the convolutional layer are set to be the ReLU function. To overcome overfitting, the

first and third convolutional layers are followed by Dropout layer whose dropout rate is 0.5.

The optimization algorithm of our proposed decoupled frequency-temporal CNN is set to be adagrad while the number of epoch and the batch size are set to be 10 and 512, respectively. The implementation of our model and the following experimental study will be released later.

B. Experimental Results

After model training, the testing accuracy with all nine features is over 99% as shown in Table II. However, the single accuracy metric is not the focus of our experimental study, considering the CWRU dataset is not a challenging one. Due to the decoupled structure of the feature and temporal modeling in the convolutional layers, the attributions from features and temporal directions could be analyzed and formed in corresponding saliency maps. From feature and temporal saliency maps, we are able to understand what are important features and time steps of data samples to predict their fault types. In the generated saliency maps, light color represents a high attention value and dark color denotes a low attention value.

TABLE II.   THE RESULTS OF UNDER DIFFERENT FEATURES COMBINATIONS

Number of Features	Testing Accuracy
9	99.4 %
6	98.5 %
4	95.8 %

As shown in Figures 3, 4, and 5, four training data samples from four classes: normal condition, inner race defect, outer race defect and ball defect that have been correctly classified by our model are randomly selected and then visualized. Grad-CAM are then applied on the first and third convolutional layer to generate feature and temporal saliency maps. These figures are able to make predictions more explainable. In Figure 3.b, it is easily shown that the model is very aggressive and leave all attention on two frequency features: spectral kurtosis and spectral skewness to predict this data sample as the normal condition. And from Figure 3.c, time steps 9 and 10 are almost non-informative to the prediction of this data sample. In the Figure 4, it could be analyzed that which part of features and time steps have received more attention for its correct prediction as the inner race fault type. As shown in Figure 4.b, the three frequency features are redundant since they are shown in dark color. The figure 4.c shows that the second half of the data sample contributes most to the model prediction. Similar analysis could be done in Figure 5 corresponding to the rest two data samples from outer race defect and ball defect. As discussed above, the generated feature and temporal saliency maps allow us to understand the impacts from features and time steps for each prediction that the model makes and obtain insights into why or how the model arrives at this outcome. With these insights, we can further improve the underlying model.

In the following, the feature saliency maps will be conducted over all training samples that are correctly classified. For each training sample, the generated feature saliency map contains the feature importance score across feature and time axis, as shown in Figure 3.b, for each specific data sample. To capture the global attention value for each feature, the entries from all saliency maps are averaged along the data axis and

---

time axis. At last, nine values corresponding to nine features could be computed in the range from zero to one which can be regarded as feature importance scores for model predictions. As shown in Figure 6, two frequency domain features: spectral skewness and kurtosis are the top-2 important features. To further validate the feature importance scores, we select the feature subset ranked by importance scores and re-train the model with the same hyper-parameter configurations. As shown in Table II, the performance was only dropped by 0.9% when the number of features is reduced from 9 to 6. And even when only three features are selected, the performance drop is still less than 4%. Therefore, it also verifies the effectiveness of the computation of feature importance scores.

Fig. 6. Features importance over all true positive training samples.
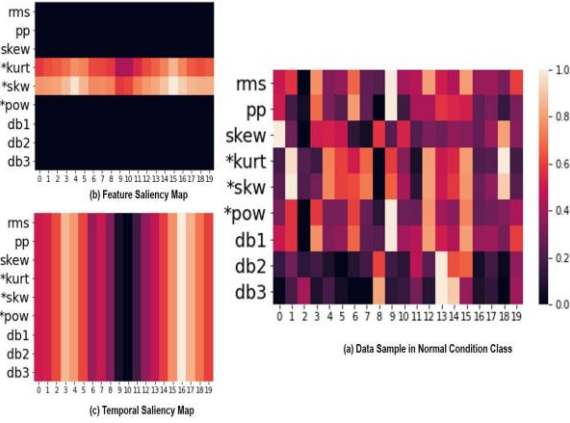
Fig. 3. Features and temporal saliency maps for a data sample who is correctly diagnosed as Normal Condition.
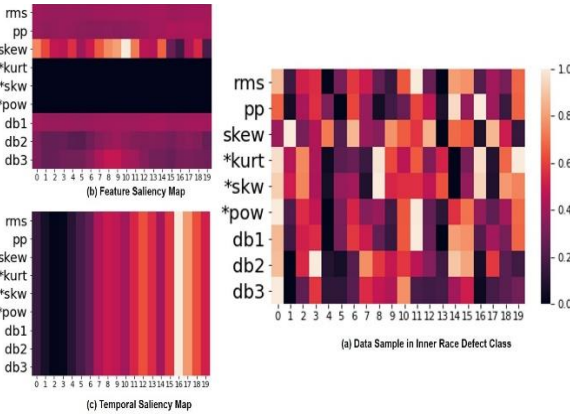
Fig. 4. Features and temporal saliency maps for a data sample who is correctly diagnosed as Inner Race Defect.
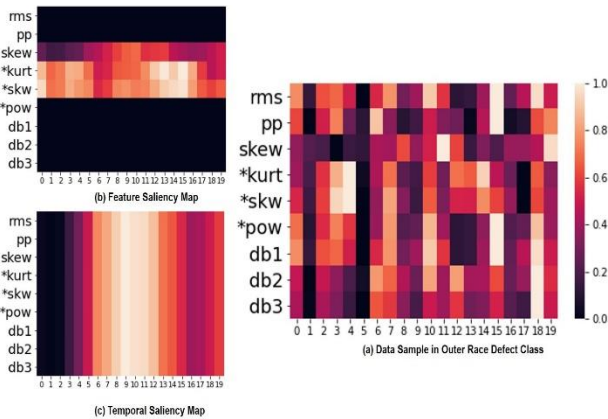
Fig. 5. Features and temporal saliency maps for a data sample who is correctly diagnosed as Outer Race Defect.

V. CONCLUSION

This paper has illustrated the lack of explainability which hinders the application of deep learning models on machine health monitoring. To crack open the black-box model, a DEFT-CNN method combined with local feature extraction was proposed. This specific structure enables the introduction of Grad-CAM to generate saliency maps explaining feature and temporal information. Experimental results indicated a good balance between precision and explainability. It also highlights that the model structure and feature selection can be further optimized and improved based on the gained insights. In the future, more experimental studies should be conducted to explore the potential edge cases. Except that, several limitations of our present framework are given as below: 1 this framework can only provide local explanation (why do you think this sample is diagnosed as normal one) instead of global one (what do you think a normal working condition sample looks like). 2 Local feature extraction is performed before DEFT-CNN. Therefore, explanations are limitations to the extracted features and sliced time windows. The above shortcomings point out our future research direction. From this paper, we hope this research direction: explainable machine health monitoring should receive more attentions from our research community.

REFERENCES

[1] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," Mechanical Systems and Signal Processing, vol. 138, p. 106587, 2020.

[2] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," Mechanical Systems and Signal Processing, vol. 115, pp. 213–237, 2019.

[3] R. X. Gao, L. Wang, M. Helu, and R. Teti, "Big data analytics for smart factories of the future," CIRP Annals, 2020.

[4] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," Mechanical Systems and Signal Processing, vol. 108, pp. 33–47, 2018.

[5] G. Helbing and M. Ritter, "Deep learning for fault detection in wind turbines," Renewable and Sustainable Energy Reviews, vol. 98, pp. 189– 198, 2018.

[6] D. Xiao, Y. Huang, C. Qin, H. Shi, and Y. Li, "Fault diagnosis of induction motors using recurrence quantification analysis and lstm with weighted bn," Shock and Vibration, vol. 2019, 2019.

[7] G. Terrazas, G. Martínez-Arellano, P. Benardos, and S. Ratchev, "Online tool wear classification during dry machining using real time cutting force measurements and a cnn approach," Journal of Manufacturing and Materials Processing, vol. 2, no. 4, p. 72, 2018.

[8] Z. Chen, M. Wu, R. Zhao, F. Guretno, R. Yan, and X. Li, "Machine remaining useful life prediction via an attention based deep learning approach," IEEE Transactions on Industrial Electronics, pp. 1–1, 2020.

[9] X. Li, W. Zhang, and Q. Ding, "Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction," Reliability Engineering & System Safety, vol. 182, pp. 208–218, 2019.

[10] J. Gertler, Fault detection and diagnosis in engineering systems. CRC press, 1998.

[11] R. Yan, R. Zhao, and R. X. Gao, "Noise-assisted data processing in measurement science: part one part 40 in a series of tutorials on instrumentation and measurement," IEEE Instrumentation & Measurement Magazine, vol. 15, no. 5, pp. 41–44, 2012.

[12] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 618–626.

[13] M. Ancona, E. Ceolini, C. Öztireli, and M. Gross, "Towards better understanding of gradient-based attribution methods for deep neural networks," arXiv preprint arXiv:1711.06104, 2017.

[14] C. Molnar, Interpretable Machine Learning. Lulu. com, 2020.

[15] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in European conference on computer vision. Springer, 2014, pp. 818–833.

[16] M. T. Ribeiro, S. Singh, and C. Guestrin, "" why should i trust you?" explaining the predictions of any classifier," in Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1135–1144.

[17] ——, "Anchors: High-precision model-agnostic explanations," in Thirty- Second AAAI Conference on Artificial Intelligence, 2018.

[18] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in neural information processing systems, 2017, pp. 4765–4774.

[19] J. Adebayo, J. Gilmer, M. Muelly, I. Goodfellow, M. Hardt, and B. Kim, "Sanity checks for saliency maps," in Advances in Neural Information Processing Systems, 2018, pp. 9505–9515.

[20] A. Shrikumar, P. Greenside, and A. Kundaje, "Learning important features through propagating activation differences," arXiv preprint arXiv:1704.02685, 2017.

[21] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, and W. Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," PloS one, vol. 10, no. 7, p. e0130140, 2015.

[22] R. Assaf and A. Schumann, "Explainable deep neural networks for multivariate time series predictions." in IJCAI, 2019, pp. 6488–6490.

[23] M. Kraus and S. Feuerriegel, "Forecasting remaining useful life: Interpretable deep learning approach via variational bayesian inferences," Decision Support Systems, vol. 125, p. 113100, 2019.

[24] N. Lee, M. H. Azarian, and M. G. Pecht, "An explainable deep learning-based prognostic model for rotating machinery," arXiv preprint arXiv:2004.13608, 2020.

[25] J. Grezmak, P. Wang, C. Sun, and R. X. Gao, "Explainable convolutional neural network for gearbox fault diagnosis," Procedia CIRP, vol. 80, pp. 476–481, 2019.

[26] R. Zhao, D. Wang, R. Yan, K. Mao, F. Shen, and J. Wang, "Machine health monitoring using local feature-based gated recurrent unit networks," IEEE Transactions on Industrial Electronics, vol. 65, no. 2, pp. 1539–1548, 2017.

[27] M. Lin, Q. Chen, and S. Yan, "Network in network," arXiv preprint arXiv:1312.4400, 2013.

[28] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the case western reserve university data: A benchmark study," Mechanical Systems and Signal Processing, vol. 64, pp. 100–131, 2015.

Deep Feature-aligned Convolutional Neural Network for Machinery Fault Diagnosis

Junbin Chen
School of Mechanical & Automotive Engineering
South China University of Technology
Guangzhou, China
201920100401@mail.scut.edu.cn

Longcan Liu
School of Mechanical & Automotive Engineering
South China University of Technology
Guangzhou, China
liulongcan0410@outlook.com

Ruyi Huang
School of Mechanical & Automotive Engineering
South China University of Technology
Guangzhou, China
201710100190@mail.scut.edu.cn

Weihua Li*
School of Mechanical & Automotive Engineering
South China University of Technology
Guangzhou, China
whlee@scut.edu.cn

Abstract—**Convolutional neural network (CNN) has been witness to remarkable development and application over the past decade in the field of fault diagnosis. However, it has an obvious limitation that the property of shift-invariance in CNN is not robust enough, resulting in insufficient robustness of feature extraction, which is manifested by fluctuant prediction accuracy and confidence. Aiming at overcoming such limitation, in this paper, a deep feature alignment method is proposed for extracting aligned features from periodical vibration signals, and applied for the fault diagnosis of rotating machinery. First, considering the periodicity of bearing vibration signals, a feature-aligned CNN (FACNN) architecture is constructed to improve the shift-invariance of CNN. Second, the performance of the proposed structure is compared with the traditional CNN in three aspects: diagnosis accuracy, prediction confidence and feature robustness. Finally, a bearing fault diagnosis case was carried out on an experimental setup of machine tool to validate the effectiveness of FACNN. Experimental results have shown the FACNN outperforms the traditional CNN in the task of fault diagnosis for rotating machinery.**

Keywords—fault diagnosis, feature-alignment, convolutional neural network, periodicity signals, rotating machinery

I. INTRODUCTION

Convolutional neural network (CNN) has demonstrated its powerful performance in image recognition [1] and fault diagnosis [2]. However, recent research has shown that the performance of shift-invariance in CNN is not robust enough, as the prediction accuracy and confidence of CNN fluctuates dramatically with small input translation. Therefore, the features extracted from the original data may vary from different offset, especially for the periodical signals [3]. With the signals being sliced or segmented isometrically, different samples obtain different translations, which will reduce the robustness of the features extracted by CNN, making it hard to have further performance breakthroughs. Thus, it is necessary to specially design the structure of CNN to fit the characteristics of periodical vibration signals, enhance its feature extraction performance and finally improve the performance of the network [4].

Generally speaking, there are three ways to extract features from vibration signals: using signal processing methods [5], learning by artificial neural networks [6], [7], and combining these two methods [8]. In the study of feature engineering, researchers mainly focus on improving the model structure or optimizing the model parameters. For example, Zhang et al. propose a specially designed CNN structure: a large size of convolution kernel is applied in the first layer of CNN to suppress high-frequency noise, and a new layer called Adaptive Batch Normalization(AdaBN) is used to reduce the sensitivity of the extracted feature to different working conditions, which further improves the domain adaptability of the network [9]. Xiao et al. use Stacked Denoising Auto-encoders(SDAE) to adaptively extract features from asynchronous motor and have achieved better performance on fault diagnosis compared to traditional motor current signature analysis(MCSA) [10]. Wu et al. construct feature combinations by Empirical Mode Decomposition(EMD) and Wavelet Packet Transform(WPT), and then use Kernel Principal Component Analysis (KPCA) to achieve feature fusion [11]. Xie et al. design a CNN structure and train the frequency spectrums of four components after a two-level discrete wavelet transform (DWT) [12]. Lu et al. design a self-normalization CNN(SCNN) to achieve faster convergent rate [13]. Wang et al. combines CNN and SDAE to reduce the influence of environment noise [14]. A new CNN structure by improving the feedback mechanism using wavelet packet transform(WPT) was proposed in [15] and experiment results have shown that the proposed method has high accuracy in fault diagnosis. It can be concluded that most fault diagnosis methods involving CNN are used in combination with other mature techniques. Few research focuses on studying the specific structure of CNN and working on how to improve the performance of CNN itself. Therefore, since we notice the limited shift-invariance of CNN, it is necessary to figure out the reason and give a reasonable solution.

Aiming at overcoming the limited shift-invariance of traditional CNN, and further extracting aligned features from periodical vibration signals, in this paper, a deep feature alignment method is proposed and applied for the fault diagnosis of rotating machinery. Considering the periodicity of bearing vibration signals, a feature-aligned CNN (FACNN) architecture is constructed by specially designing the hyper-parameters of the convolutional and max-pooling layer, which improves the shift-invariance of CNN. Then, to fully explain the performance of the proposed structure, FACNN is compared with a traditional CNN structure in three different aspects: diagnosis accuracy, prediction confidence and feature robustness. Finally, a bearing fault diagnosis case was carried out on an experimental setup of machine tool to validate the effectiveness of FACNN.

---

*Weihua Li is the corresponding author. (e-mail: whlee@scut.edu.cn).

The main contributions of our research are concluded as follows:

- A novel diagnosis model named feature-aligned CNN (FACNN) is constructed to improve the shift-invariance of CNN, which can adaptively extract

  aligned features from periodical vibration signals and further improve the accuracy and confidence of fault diagnosis.

- Full-period max-pooling and single-stride convolution are introduced into the FACNN to keep features extracted by max-pooling layer aligned and to make the convolutional layer shift-equivariant, which enhance the robustness of features extracted from different samples and improves the generalization of the proposed FACNN.

The subsequent parts of this paper are organized as follows. Section II introduces the basic theory of shift-invariance in traditional CNN and analyzes its limitations. The specific procedure of FACNN are presented in Section III, followed by a bearing fault diagnosis case and experimental validation in Section IV. The final section draws a conclusion and discusses some future perspectives of our work.

II. BASIC THEORY OF SHIFT-INVARIANCE IN TRADITIONAL CNN

A. Shift-equivariance and Shift-invariance

A function $f(\bullet)$ is defined to be shift-equivariant, if a certain translation is applied to its input and the same translation occurs to its output. The shift-equivariance of a function can be described as follows:

$$f\left(shift\left(x\right)\right) = shift\left(f\left(x\right)\right) \tag{1}$$

Similarly, a function is defined to be shift-invariant, if a certain translation is applied to its input, and nothing occurs to its output. The shift-invariance of a function can be described as follows:

$$f\left(shift\left(x\right)\right) = f\left(x\right) \tag{2}$$

B. Foundation of CNN's shift-invariance

The traditional structure of CNN is usually assumed to be shift-invariant, which means that the same object would be correctly detected even though a translation is applied to the input. The shift-invariance of CNN is actually contributed by the shift-equivariance of convolutional layer and the shift-invariance of max-pooling layer, which is shown in Fig. 1. On the one hand, the shift-equivariance of convolutional layer is
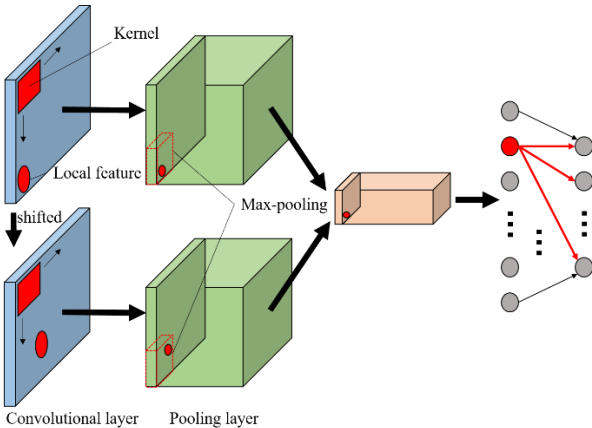
simple to explain, because the kernel, which can be seen as a feature detector, will not detect the object(or feature) until it moves to the corresponding right position and matches the object. On the other hand, max-pooling layer is considered to be shift-invariant, as long as the largest value is still within its receptive field(input to the pooling function) after translation, which leads to the same result of max-pooling. The combination of convolution layer and max-pooling layer makes the output of CNN insensitive to a certain small translation of input.

Considering one-dimensional CNN, the principle and conditions of its shift-invariance are fully discussed as follows.

1) Shift-equivariance of convolutional layer

In a traditional CNN, the convolutional layer is one of the two most commonly used layers to extract effective features from raw vibration signals, which can be expressed as follow:

$$z = Convolution\left(x\right) = \sigma\left(K \otimes x + b\right) \tag{3}$$

where x and z represent the input and output of the convolutional layer, respectively, K and b represent the convolution kernel weight and bias term, $\otimes$ means the convolutional operation between K and x, and $\sigma(\bullet)$ is an activation function, which is used to establish a nonlinear mapping between x and z.

The principle of the convolutional layer to be shift-equivariant is shown in Fig. 2. Suppose that kernel K has a higher output for the subsequence $[a_1, a_2, a_3]$ in the input, that is, the kernel can be used to detect this subsequence. It can be seen from Fig. 2 that $[a_1, a_2, a_3]$ was originally located at position L. When the input is shifted by ΔL, if and only if ΔL is an integer multiple of the convolution stride s that kernel K can still exactly correspond to $[a_1, a_2, a_3]$, which means that the convolutional operation before and after translation only changes the position of the output, that is, shift-equivariance is achieved. Therefore, the condition for the convolutional layer to achieve shift-equivariance is:

$$\Delta L = ns, \quad n \in Z \tag{4}$$

The position difference Δl between the new and original output is exactly equal to the ratio of the input translation and the stride of the convolution kernel. If Eq. (4) cannot be satisfied, kernel K will not exactly correspond to $[a_1, a_2, a_3]$, resulting in the value of new output different from the original one, which finally makes the extracted features distorted.

2) Shift-invariance of max-pooling layer

Another commonly used layer in a traditional CNN is pooling layer, which can be expressed as follow:

$$h = Pooling\left(z\right) \tag{5}$$

Fig. 2. Shift-equivariance of convolutional layer

Fig. 1. Illustration of shift-invariance in traditional CNN
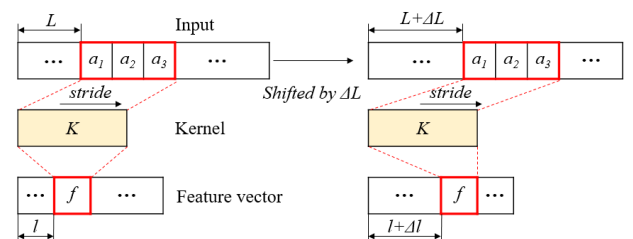
where z and h are the output feature map of the convolutional layer and pooling layer, respectively. In a pooling operation, max-pooling is proved to be one of the most powerful pooling techniques to extract effective features from the input feature vector.

The principle of the max-pooling layer to be shift-invariant is shown in Fig. 3. Similar to the convolutional layer, when the input translation of the max-pooling layer is an integer multiple of the stride, the max-pooling operation is shift-equivariant, and the output translation is still the ratio of the input translation and the value of stride. When the input translation is less than the value of stride, the max-pooling layer is shift-invariant, if and only if the translation does not cause any change in the maximum value of any pooling window. As shown in Fig. 3, the max-pooling layer is shift-invariant, when the maximum value in the blue pooling window remains m_1 and that in the green pooling window remains n_2 after translation, and so on. In other cases, the max-pooling layer is neither shift-equivariant nor shift-invariant, and the pooled output is also distorted. The smaller the input translation or the longer the length of the pooling window is, the greater probability that the maximum value in the pooling window does not change after translation, which means the greater probability that the max-pooling layer is shift-invariant.

C. Limitation of CNN's shift-invariance

From the former discussion of shift-equivariance and shift-invariance, it can be concluded that the property of shift-invariance of traditional CNN suffers from the limited shift-equivariance of the convolutional layer and the limited shift-invariance of the max pooling layer. To make CNN strictly shift-invariant for effective features, the convolutional layer and max-pooling layer need to be shift-equivariant and shift-invariant respectively. However, in bearing fault diagnosis, since the initial phase of the signal is not locked when the signal is collected, and the period of fault features is unknown in advance, the random isometric dividing of samples leads to different initial phases of different samples, which in turn makes it impossible to determine the translation of input. In this case, due to the limited shift-invariance of traditional CNN, local features with position variance in different samples may show different degrees of distortion after passing through a certain convolution layer, and the probability and degree of feature distortion will keep increasing in the subsequent process. Therefore, to solve the problem of distorted features due to input translation, and to improve the effectiveness of the extracted features, it is necessary to carefully consider the specific design of CNN.

III. PROPOSED METHODS

It can be concluded from Section II that the shift-invariance of CNN is actually influenced by the hyper-parameters of the network, such as the receptive field and the stride. Therefore, to improve the shift-invariance of traditional CNN, as well as considering the periodicity of vibration signal, Feature-aligned Convolutional Neural Network(FACNN) is
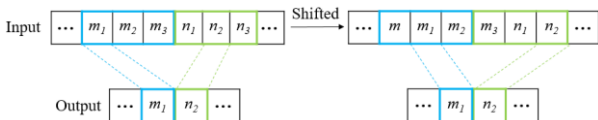
proposed. The FACNN does not change the internal structure of a traditional CNN, but its hyper-parameters are well designed to best deal with vibration signal. The structure of FACNN is shown in Fig. 4.

A. Single-stride Convolution

According to the above discussion on shift-equivariance, when the sample translation is just an integer multiple of the convolution stride, the translation equivalence between output and input can be guaranteed, that is, the activation value of output features will not be distorted due to the translation of input. Due to the randomness of the translation of input, the convolution stride is set to be 1, which can divide any value of translation. The single-stride convolution can not only ensure that the extracted features are not distorted, but also more information of the input can be obtained. The working principle of single-stride convolution is shown in Fig. 5.

Suppose that the original input feature vector of the network is $[A, B, C, D, A, B, C, D]$, and the new input is obtained by shifting the original input to the left by one bit. Both inputs are convolved by the same kernel, whose mapping is $f(\bullet)$.

1) Single-stride

When the convolution stride is 1, which must divide the value of translation. The output feature vector corresponding to the input before translation is:

$$f_o = Convolution(\boldsymbol{x}) \\ = [f_{AB}, f_{BC}, f_{CD}, f_{DA}, f_{AB}, f_{BC}, f_{CD}] \tag{6}$$

After translation, the corresponding output is:

$$f_o^{shifted} = Convolution(\boldsymbol{x}^{shifted}) \\ = [f_{BC}, f_{CD}, f_{DA}, f_{AB}, f_{BC}, f_{CD}, f_{DA}] \tag{7}$$

It can be seen that the output is also shifted to the left by one bit, that is, the convolutional layer is shift-equivariant.

2) Multiple-stride

When the convolution stride is 2, which cannot divide the value of translation. The output feature vector corresponding to the input before translation is:

$$f_o = Convolution(\boldsymbol{x}) = [f_{AB}, f_{CD}, f_{AB}, f_{CD}] \tag{8}$$

After translation, the corresponding output is:

Fig. 3. Shift-invariance of max-pooling layer
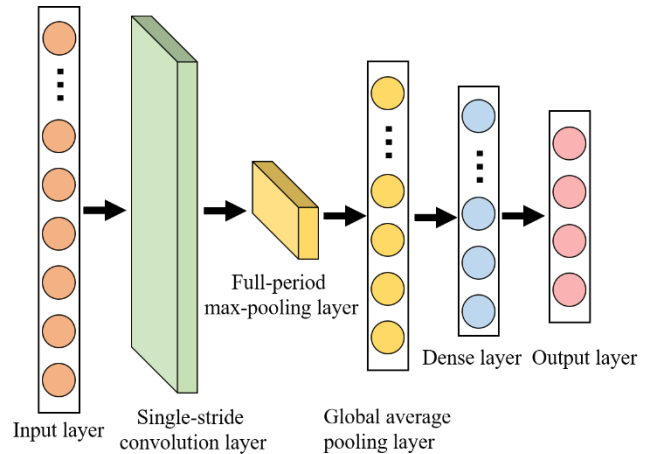
Fig. 4. Structure of FACNN
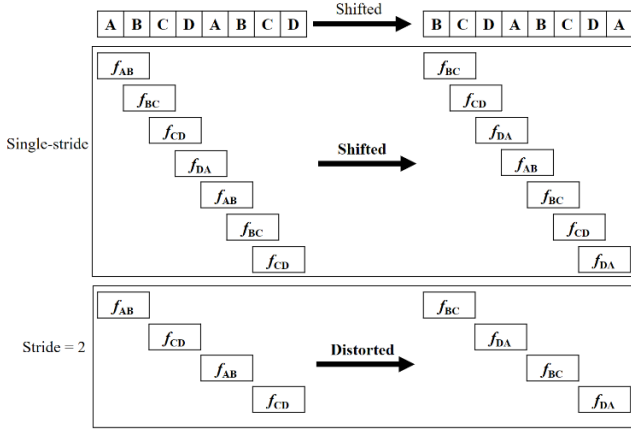
Fig. 5. Single-stride convolution.

$$f_o^{shifted} = Convolution\left(x^{shifted}\right) = \left[f_{BC}, f_{DA}, f_{BC}, f_{DA}\right] \quad (9)$$

The features of the new output are different from the original output, that is, the extracted features are distorted.

B. Full-period Max-pooling

The local fault features extracted by the single-stride convolutional layer still have different positions between different samples, which is likely to cause distortion in subsequent operations. Therefore, a full-period max-pooling layer is applied to align the features extracted by the convolutional layer at once. As mentioned earlier, vibration signals are usually periodic and thus fault features are also periodic. When the size of the pooling window is larger than the period of the fault feature cycle, then each pooling window can capture at least one effective local feature, which guarantees feature alignment between different samples. The principle that the proposed max-pooling layer achieve feature alignment is shown in Fig. 6. It shows the pooling process of the original feature vector and the shifted feature vector under different sizes of pooling window. On the left, a green pooling window, whose size is less than the length of effective feature period is used, indicating that the effective features between the two output feature vectors remain position differences. On the right, a blue pooling window, whose size is equal to the length of effective feature period is used, indicating that the effective features between the two output feature vectors are aligned.

In fact, the period of effective fault features is unknown in advance, but the impact period of bearing faults is usually less than the shaft rotation frequency. Therefore, the max-pooling window and stride are set to be the length of the shaft rotation period.
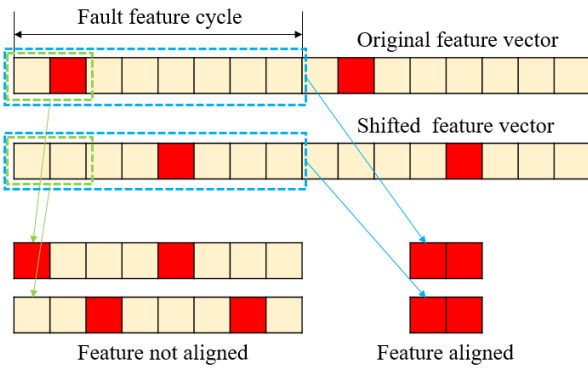
Fig. 6. Full-period max-pooling

C. Global Average Pooling

Due to the periodicity of vibration signals, there are usually multiple periods in a sample, which makes the feature vectors redundant. To obtain a more robust expression of the features, global average pooling (GAP) is applied to each feature vector, which is shown in Fig. 7.

The left side of the figure shows the process of dimension reduction of the output feature matrix using a traditional flatten layer. It directly combines the feature vectors of multiple channels as output, resulting in a large output dimension and redundant features. The right side of the figure applies the GAP layer to reduce the dimension of the feature matrix, which not only reduces the size of the output, but also effectively reduce the influence of abnormal feature values on network output.

Based on the above discussion, the architechture and hyper-parameters of FACNN are determined, as shown in TABLE I. As a comparison, a traditional CNN architechture is constructed, which isshown in TABLE II.

IV. EXPERIMENTAL VALIDATION

To validate the effectiveness of FACNN, a bearing fault diagnosis experiment is carried out on an experimental setup of machine tool. The experiment involves four types of bearing health state: normal, inner race defect, outer race defect, and cage defect. For each bearing health state, the vibration signal is collected under three different loads, which are simulated by idling(unload), stainless steel cutting and aluminum cutting, respectively. The specific experiment setting is shown in TABLE III.

Four diagnosis cases are set up, which is shown in TABLE IV. For each training condition, the training set is composed of 1000 samples from each bearing state. The performance of FACNN and traditional CNN will be compared in detail from three aspects: diagnosis accuracy, prediction confidence of the correct diagnosis sample, and deep feature alignment extraction.

A. Comparison of Diagnosis Accuracy

Each experiment is repeated five trials to obtain the average diagnosis accuracy of FACNN on the test set, which is compared with traditional CNN, as shown in Fig. 8 and TABLE V.

It can be seen that the diagnosis accuracies of FACNN in the four cases are higher than that of traditional CNN. Besides, the diagnosis accuracy of traditional CNN has dropped sharply in case 3, while FACNN maintains excellent

Fig. 7. GAP

289
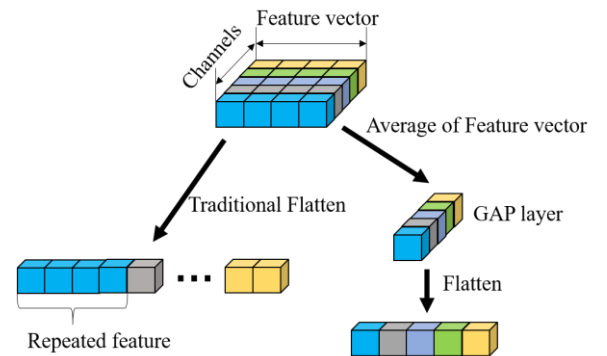
TABLE I. EXPERIMENTAL STRUCTURE OF FACNN

Type of Layer	Parameters	Values	Activation Function	Output shape
Input	/	/	/	(2048, 1)
Conv1d	Size/Stride/Padding	65×64/1/Same	ReLU	(2048, 64)
Max-Pooling	Size/Stride/Padding	256/256/Valid	/	(8, 64)
GAP	/	/	/	(64)
Dense	Number of Nodes	32	ReLU	(32)
Output	/	/	SoftMax	(4)

TABLE II. STRUCTURE OF TRADITIONAL CNN

Type of Layer	Parameters	Values	Activation Function	Output shape
Input	/	/	/	(2048, 1)
Conv1d_1	Size/Stride/Padding	65×8/8/Same	ReLU	(258, 8)
Max-Pooling_1	Size/Stride/Padding	8/8/Valid	/	(32, 8)
Conv1d_2	Size/Stride/Padding	9×16/4/Same	ReLU	(8, 16)
Max-Pooling_2	Size/Stride/Padding	2/2/Valid	/	(4, 16)
Flatten	/	/	/	(64)
Dense	Number of Node	32	ReLU	(32)
Output	/	/	SoftMax	(4)

TABLE III. EXPERIMENT SETTING

Condition	8000rpm
Unload	S1
Aluminum Cutting	T
Stainless Steel Cutting	S2

TABLE IV. DIAGNOSIS CASES SETTING

Case No.	Training Set	Test Set
Case 1	S1, S2, T	T
Case 2	S1, S2	T
Case 3	S1	T
Case 4	S2	T

Fig. 8. Diagnosis accuracy of traditional CNN and FACNN

TABLE V. DIAGNOSIS ACCURACY OF TRADITIONAL CNN AND FACNN

Methods	Case 1	Case 2	Case 3	Case 4
Traditional CNN	99.93	99.94	98.41	99.91
FACNN	100.00	100.00	99.98	100.00
Improvement	**0.07**	**0.06**	**1.57**	**0.09**

(a) Case 1   (b) Case 3

Fig. 9. Prediction confidence of traditional CNN and FACNN

TABLE VI. PREDICTION CONFIDENCE OF TRADITIONAL CNN AND FACNN

		Case 1	Case 3
Mean	Traditional CNN	1.000	0.983
	FACNN	1.000	**0.994**
Standard deviation	Traditional CNN	$2.377×10^{-4}$	$3.409×10^{-2}$
	FACNN	$\mathbf{4.721×10^{-5}}$	$\mathbf{5.085×10^{-3}}$

performance, with a 1.57% accuracy improvement, indicating that FACNN has better generalization than traditional CNN.

B. Prediction Confidence

Prediction confidence is an indicator from zero to one, referring to the output probability of a network for a certain category, which represents the confidence degree of a network for this category. Prediction confidence analysis of the correctly diagnosed samples can indicate the robustness of the features extracted by the network. Here, case 1 and 3 are used for inspection. 200 samples that are both correctly classified by traditional CNN and FACNN from each case are selected randomly. Sample No.1 to No.50 are normal samples. Sample No.51 to No.100 are inner race fault samples. Sample No.101 to No.150 are outer race fault samples; and sample No. 151 to No.200 are cage fault samples.

The results are shown in Fig. 9 and TABLE VI. For samples from each bearing health state, the mean and standard deviation of their prediction confidence are calculated, and the statistical indicators under various health states are averaged. It can be concluded from the results that the proposed FACNN has higher and more stable prediction confidence of correctly diagnosed samples than traditional CNN, which indirectly proves that the features extracted by FACNN are more robust.

C. Analysis of Deep Feature alignment Extraction

In this part, the output of the fully-connected layer of FACNN and traditional CNN are used to analyze the consistency of features extracted by the network from the samples that are under the same bearing state and working condition. Similarly, case 1 and case 3 are investigated. Normal samples from the test set are used as input, from which 100 samples are randomly selected to observe the output of the feature layer of the two networks. First, for each health state under target working condition, the standard deviation between these 100 samples of each feature is calculated. Then, the standard deviation of the various features and various bearing states are averaged as the indicator to analyze the degree of feature alignment. The results are shown in Fig. 10 and TABLE VII.

It can be seen from Fig. 10 that in the foregoing two cases, the features extracted by traditional CNN show obvious differences, indicating that traditional CNN is not stable enough to extract robust features. Besides, the average standard deviation of the features extracted by FACNN in each case is much smaller than that extracted by traditional CNN, indicating that FACNN has much higher consistency than traditional CNN does in extracting features, that is, the features extracted by FACNN are more robust, as the features of different samples are aligned.

a) Traditional CNN_Case 1   b) Traditional CNN_Case 3

c) FACNN_Case 1   d) FACNN_Case 3
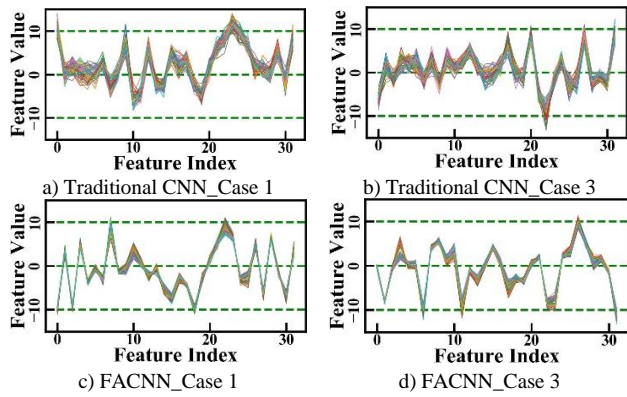
Fig. 10. Output of feature layers of traditional CNN and FACNN

TABLE VII.   FEATURES EXTRACTED BY TRADITIONAL CNN AND
FACNN

		Case 1	Case 3
Standard deviation	Traditional CNN	1.202	1.188
	FACNN	**0.590**	**0.626**

V. CONCLUSION

For rotating machinery fault diagnosis, the finite shift-invariance of traditional CNN limits its further performance. Therefore, in this paper, FACNN is proposed to solve this problem. In the proposed feature-aligned network, several techniques, such as single-stride convolution, full-period max-pooling and global average pooling are applied to best fit the characteristics of vibration signals and make the CNN shift-equivariant again. Experiments have shown that the proposed FACNN does obtain better performance in accuracy, prediction confidence and feature alignment.

A future direction of FACNN is to combine signal processing to enhance the interpretability of the extracted features, which will help us better understand how they actually affect the generalization of the network. Another possible direction is to combine transfer learning to further improve the generalization of the network, making it more adaptable to more complex transfer tasks.

REFERENCES

[1] Joseph Redmon and Ali Farhadi, "YOLOv3: An Incremental Improvement" arXiv:1804.02767v1 [cs.CV], Apr 2018

[2] R. Huang, Y. Liao, S. Zhang and W. Li, "Deep Decoupling Convolutional Neural Network for Intelligent Compound Fault Diagnosis," IEEE Access, vol. 7, pp. 1848-1858, 2019.

[3] Aharon Azulay and Yair Weiss, "Why do deep convolutional networks generalize so poorly to small image transformations?" Journal of Machine Learning Research 20 (2019) 1-25.

[4] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R. X. Gao, "Deep learning and its applications to machine health monitoring," Mech. Syst. Signal Process., vol. 115, pp.213-237, Jan. 2019.

[5] W. Li, S. Zhang, S. Rakheja. "Feature denoising and nearest–farthest distance preserving projection for machine fault diagnosis," IEEE Transactions on Industrial Informatics, vol. 12, no. 1, pp. 393-404, February 2016.

[6] H. Schulza, K. Chob, T. Raikob, and S. Behnkea, "Two-layer contractive encodings for learning stable nonlinear features," Neural Network, vol. 64, pp. 4-11, April 2015.

[7] R. Huang, J. Li, W. Li and L. Cui, "Deep Ensemble Capsule Network for Intelligent Compound Fault Diagnosis Using Multisensory Data," IEEE Trans. Instrum. Meas., vol. 69, no. 5, pp. 2304-2314, May 2020, doi: 10.1109/TIM.2019.2958010

[8] Z. Chen and W. Li, "Multisensor Feature Fusion for Bearing Fault Diagnosis Using Sparse Autoencoder and Deep Belief Network," in IEEE Transactions on Instrumentation and Measurement, vol. 66, no. 7, pp. 1693-1702, July 2017, doi: 10.1109/TIM.2017.2669947.

[9] W. Zhang, G. Peng, C. Li, Y. Chen and Z. Zhang, "A new deep learning model for fault diagnosis with good Anti-Noise and domain adaptation ability on raw vibration signals," Sensors, vol. 17, no. 2, February 2017.

[10] N. Xiao et al., "Adaptive feature extraction based on Stacked Denoising Auto-encoders for asynchronous motor fault diagnosis," 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Datong, 2016, pp. 854-859, doi: 10.1109/CISP-BMEI.2016.7852830.

[11] G. Wu, Y. Tian and W. Mao, "On line fault diagnosis of rolling bearings based on feature fusion," 2018 Chinese Control And Decision Conference (CCDC), Shenyang, 2018, pp. 2769-2774, doi: 10.1109/CCDC.2018.8407596.

[12] Y. Xie and T. Zhang, "Feature extraction based on DWT and CNN for rotating machinery fault diagnosis," 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, 2017, pp. 3861-3866, doi: 10.1109/CCDC.2017.7979176.

[13] K. Lu, T. Lin, J. Xue, J. Shang and C. Ni, "An Automated Bearing Fault Diagnosis Using a Self-Normalizing Convolutional Neural Network," 2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE), Zhangjiajie, China, 2019, pp. 908-912, doi: 10.1109/QR2MSE46217.2019.9021151.

[14] Y. Wang, M. Han and W. Liu, "Rolling Bearing Fault Diagnosis Method Based on Stacked Denoising Autoencoder and Convolutional Neural Network," 2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE), Zhangjiajie, China, 2019, pp. 833-838, doi: 10.1109/QR2MSE46217.2019.9021126.

[15] F. An, "Rolling Bearing Fault Diagnosis Algorithm Based on FMCNN-Sparse Representation," in IEEE Access, vol. 7, pp. 102249-102263, 2019, doi: 10.1109/ACCESS.2019.2931616.

Research on Defects Recognition Method Based on Impedance Information

Xu Zhang
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
18380155635@163.com

Jian Zhang
Shanghai Space Propulsion Technology
Research Institute
Shanghai, China
zjandy86@126.com

Libing Bai*
School of Automation Engineering
University of Electronic Science and
Technology of China
bailb991@163.com

Lulu Tian
School of Automation Engineering
University of Electronic Science and
Technology of China
tl.110119@163.com

Jie Zhang
School of Automation Engineering
University of Electronic Science and
Technology of China
zhj06_19@uestc.edu.cn

Yuhua Cheng
School of Automation Engineering
University of Electronic Science and
Technology of China
yhcheng@uestc.edu.cn

Abstract—**Eddy current pulsed thermography (ECPT) can acquire the transient thermal image of conductive materials. The defects profile obtained by thermal image is not clear due to the thermal diffusion effect. There is a lot of information in the thermal image that's not being utilized. The distribution of eddy current can be obtained from the thermal image. This paper proposes a method to extract the distribution of impedance from the distribution of eddy current through electric field constraints. The distribution of impedance can recognize the defects profile more clearly. Experimental results show that the defects profile recognition using the distribution of impedance is more accurate than the others.**

Keywords—distribution of impedance, eddy current pulsed thermography, distribution of eddy current, electric field constraints, defects profile

I. INTRODUCTION

Eddy current pulsed thermography (ECPT) has been applied for a wide range of conductive materials. It extracts useful information from the electrical, magnetic, thermal characteristics. This technology is widely used in aerospace, nuclear power, transportation and other fields. It is an essential part of the inspection process of machinery equipment. The advantage of this technology is that it can perform accurate non-contact detection of large-scale and long-distance materials in a short time [1].

The thermography is applicable to a wide range of materials, including glass fiber reinforced polymer specimen, carbon fiber reinforced polymer composites [2], power electronic devices [3] and steel [4] with great success. B. Weekes estimated the probability of detection (POD) of fatigue cracks in steel, titanium and nickel-based super alloy [5]. J. Wilson detected multiple cracks from rolling contact fatigue in rail track in a single measurement [6]. The wealth of information of ECPT transient pattern has attracted a wide interest. Several transient response features have been used for quantification of defect, which is critical for acceptance/rejection decisions for maintenance and lifetime prediction. Y. He compared two detections modes (transmission mode and reflection mode) for wall thinning and inner defects characterisation using time to peak [7]. Smyl, D proposed a detection and reconstruction method of complex structural cracking patterns with electrical imaging.

The progressive crack in rectangular beam geometry is simulated numerically and the influence of image reconstruction method on the crack estimation is also discussed [8]. Dong Liu proposed a new shape reconstruction framework rooted in the concept of Boolean operations for electrical impedance tomography. This method is tolerant to modeling errors caused by background inhomogeneity and is also quite robust to the selection of control points [9]. The electrical impedance tomography is a medical imaging technique, which estimates the distribution of impedance within a body from boundary measurements. Nowadays it is broadly applicable in the field of non-destructive testing. Alessandro Cultrera present the first electrical resistance tomography (ERT) measurements on large-area graphene samples. They characterised the electrical conductivity of chemical-vapour deposited graphene samples by performing ERT [10]. The traditional method focuses on the infrared image feature extraction. Bai, L employed a single channel blind source separation algorithm to find anomalous patterns from the transient thermal pattern. The method enables spatial and time patterns to be extracted according to the whole transient response behavior [11]. Zhu, P proposed a method which include entropy-based image selection, local sparse and low-rank decomposition to increase the contrast of defect area and background. The method reserved more meaningful defect information of experimental samples from raw ECPT data [12]. Cheng, Yuhua utilized the independent component analysis (ICA) to remove the influence of weld inhomogeneity and enhance the defect information in the stainless steel weld defect detection [13].

In this paper, a new defects recognition method based on impedance is proposed to recognize the defects profile. In this method, the heat distribution of the infrared image is transformed into eddy current distribution. The surface of the conductor sample is equivalent to a resistance network. When the coil is parallel to the surface of the conductor sample, the electric field intensity, induced by high-current electromagnetic pulses in the coil, is approximately equal in each row. The electric field intensity can be obtained from the eddy current distribution in non-defect area. At last, the eddy current distribution is transformed into impedance distribution.

This paper is organized as follows. Section II discusses the experimental process. Section III introduce the

---

methodology of the proposed method. And Section IV concludes the work in this paper.

II. EXPERIMENTAL PROCESS

During the test phase, high-current electromagnetic pulses in the coil will induce eddy currents on the surface of the conductive sample, as shown in Fig. 1. The infrared camera records the heat distribution from Joule heating and heat diffusion procedure. The heat distribution is transformed into eddy current distribution through (1). Eddy current distribution is transformed into impedance distribution through electric field confinement. The impedance distribution can recognize the location and profile of defects more clearly. The focus of this paper is the conversion of eddy current distribution into impedance distribution. The experimental process is shown in Fig. 2.
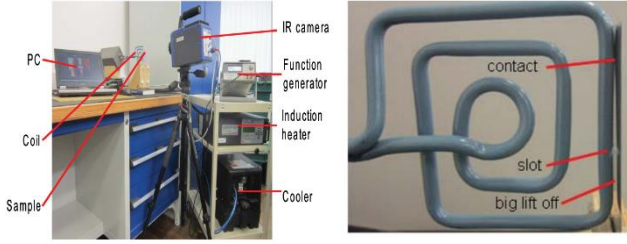
Fig. 1. Experiment set-up
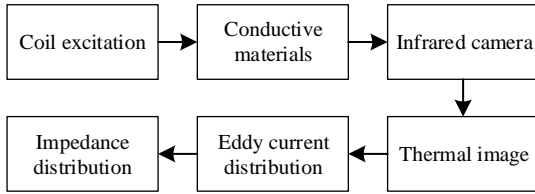
Fig. 2. Schematic diagram of the experimental process

III. THEORETICAL CONSIDERATIONS

On the surface of the conductive sample, when the eddy current encounters discontinuity of the impedance, it's flow direction and the temperature distribution will change. The eddy current will take the path of least impedance on the surface and bypass the area of high impedance. When the eddy current encounters a defect with large impedance, it will bypass the defect and form high temperature areas at the left and right ends of a defect. What's more, two low-temperature areas will be formed in the upper and lower parts of the defect, as shown in Fig. 3. At the abscissa of 50 and 300 there are two vertical high temperature areas, which are the edges of the conductor sample. The traditional method is to find the high temperature areas on the left and right sides of the defect through thermal images to recognize the defect, but the position and profile of the defect cannot be determined.
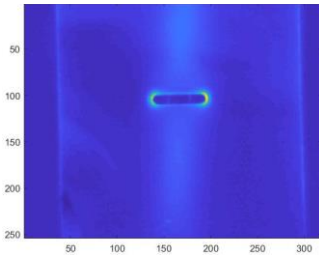
Fig. 3. The transient thermal image

The eddy current density can be obtained by temperature

$$\frac{\partial T}{\partial t} = \frac{\frac{1}{\sigma}|J_s|^2 t}{\rho C_p} \qquad (1)$$

where J_s is the eddy current density, T is the temperature, ρ is the density of sample, C_p is the specific heat capacity. According to Kirchhoff's current law, the eddy current direction can be determined by the iterative algorithm, as shown in Fig. 4. Eddy currents are induced by ac signals, so the direction of eddy currents is transient. At the left and right ends of the defect, the eddy current bypasses the defect and forms high current areas. The high current forms high temperature areas. However, the eddy current distribution is still difficult to determine the profile of the defect, so the eddy current distribution needs to be converted into the impedance distribution.
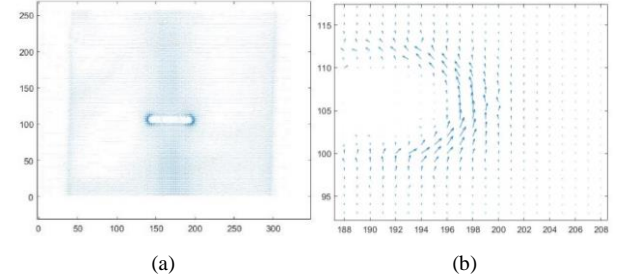
Fig. 4. The eddy current distribution (a) Global. (b) Partial.
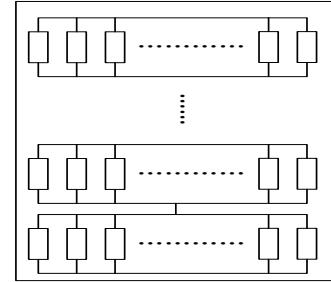
Fig. 5. Equivalent resistance network

The impedance of the sample surface is approximately the reciprocal of the electrical conductivity, so the impedance is determined by the electric field intensity and the eddy current density.

$$Z = 1/\sigma = E/J_s \qquad (2)$$

where Z is the impedance of the conductive sample surface, E is the electric field intensity of the conductive sample surface, J_s is the eddy current density of the conductive sample surface, σ is the electrical conductivity of the conductive sample surface.

The surface of the conductor sample can be equivalent to a resistance network of the parallel connection of each row of resistances, as shown in Fig. 5. When the coil is parallel to the surface of the conductor sample, the electric field intensity, induced by high-current electromagnetic pulses in the coil, is approximately equal in each row. The eddy current distribution can be equivalent to the current in each row constrained by the electric field intensity. If the electric field intensity of each row can be obtained, the impedance distribution can be obtained. The electric field intensity is induced by high-current electromagnetic pulses

in the coil for a short period (typically less than 1s). The electric field intensity is obtained from the magnetic field intensity

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} = -\mu \frac{\partial \boldsymbol{H}}{\partial t} \qquad (3)$$

where $\boldsymbol{B}$ is the magnetic flux density, $\boldsymbol{H}$ is the magnetic field intensity. The magnetic field intensity is mainly induced by the current density of the coil

$$\nabla \times \boldsymbol{H} = \boldsymbol{J_c} \qquad (4)$$

where $\boldsymbol{J_c}$ is the current density of the coil. $\boldsymbol{J_c}$ is evenly distributed in the coil, so the magnetic field intensity is evenly distributed on the surface.

(a)

(b)
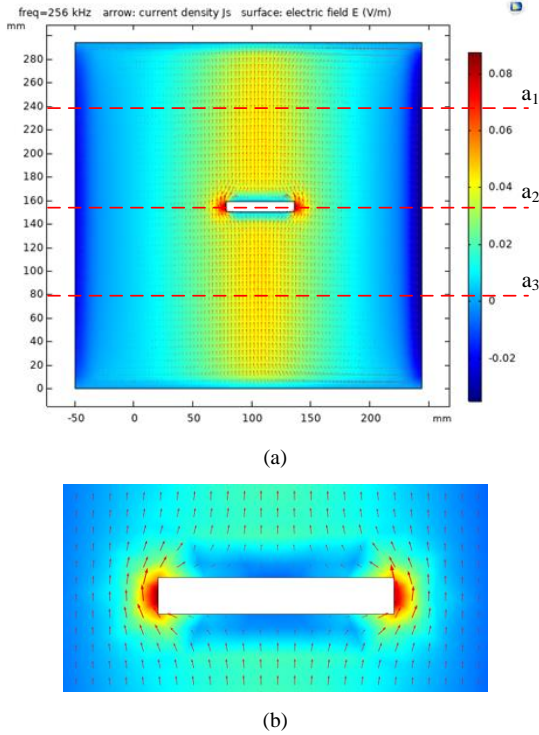
Fig. 6. The electric field distribution model: arrow: current density surface: electric field (a) Global. (b) Partial.

The electric field intensity, induced by $\boldsymbol{H}$, is evenly distributed on the surface as well. The electric field distribution model of surface can be simulated by COMSOL, as shown in Fig. 6. The electric field intensity distribution of three rows was extracted from the electric field distribution model, which is a_1 a_2 a_3. The electric field distribution in the non-defect area is a_1 and a_3. The electric field distribution in the defect area is a_2, as shown in Fig. 7.

In the electric field distribution model, the electric field intensity of each row is approximately equal in the non-defect area. In the defect area, the electric field intensity near the defect is higher. The discontinuity of impedance causes dramatic increase of electric field intensity and eddy currents. Assuming that the electric field intensity distribution in the defect and non-defect area is the same, the drastic change in current can be reflected in the impedance through (2). So the electric field intensity distribution in the non-defect area can be selected as the reference electric field. The impedance distribution can be obtained through the reference electric field and the eddy current through (2).

(a_1)

(a_2)

(a_3)
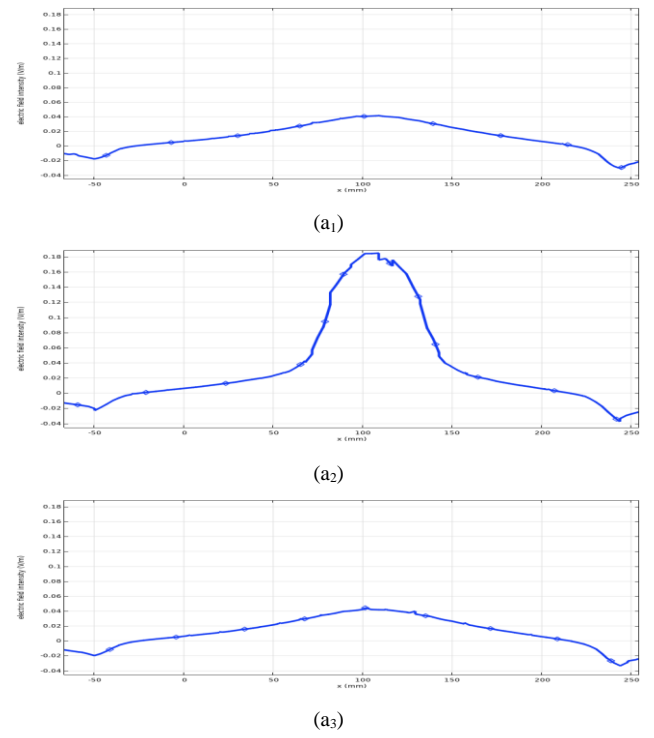
Fig. 7. Electric field density distribution of (a_1) (a_2) (a_3)

(a)

(b)

(c)
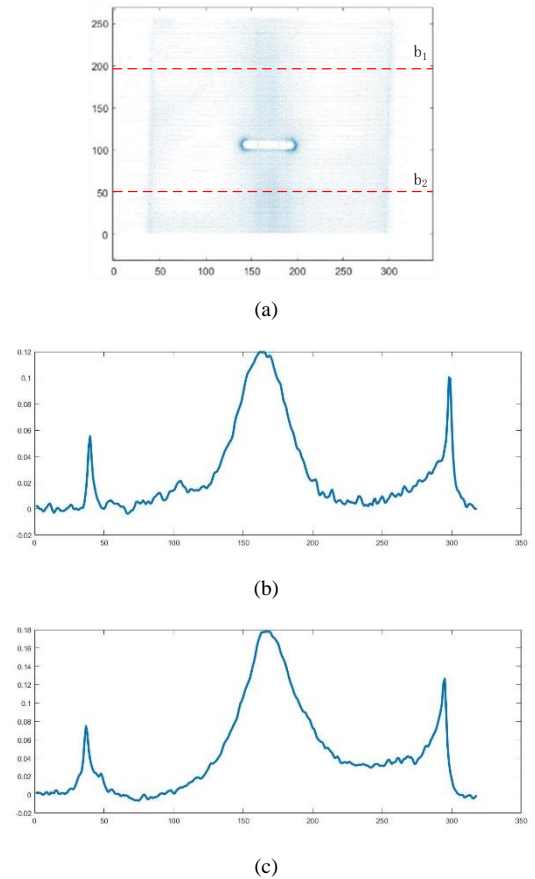
Fig. 8. (a) Test data of eddy current distribution. (b) Electric field density distribution of b_1. (c) Electric field density distribution of b_2.

The electric field intensity distribution of the two rows are extracted from test data of eddy current distribution in non-defect area, as shown in Fig. 8. In the non-defect area, the electric field intensity of each row is approximately equal. In the electric field intensity distribution, the spikes

on the left and right are the boundaries of the sample. It is caused by the edge effect of eddy current.

IV. RESULTS AND DISCUSSION

The electric field intensity in the non-defect area is obtained from the eddy current density and impedance through (2). The mean value of electric field intensity of all rows in the non-defect was calculated as the reference electric field intensity. The impedance distribution can be obtained by the reference electric field and the eddy current, as shown in Fig. 9.

(a)   (b)
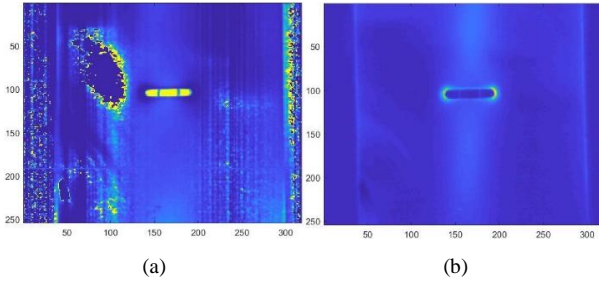
Fig. 9. (a) The impedance distribution image. (b) The thermal image.

(a)   (b)
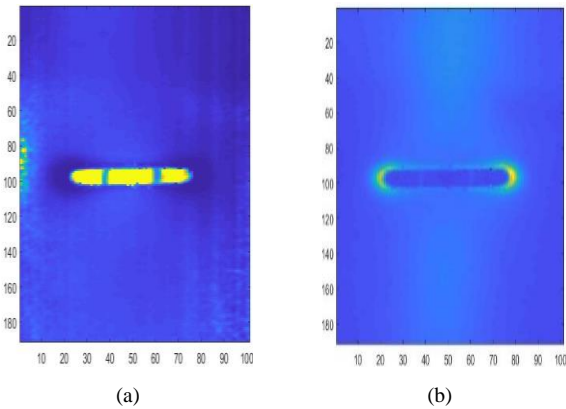
Fig. 10. (a) The impedance distribution image. (b) The thermal image.

(a)   (b)
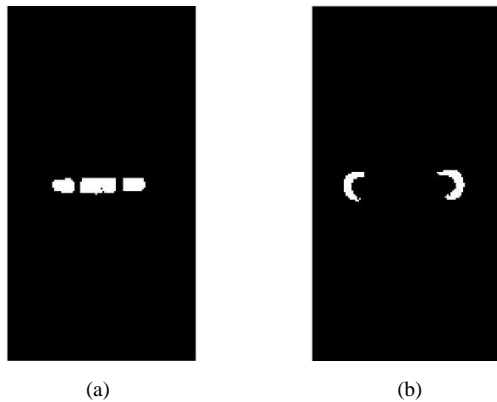
Fig. 11. The binarization image: (a) The impedance distribution image. (b) The thermal image.

The defect profile in the impedance image is clearer than that in the thermal image. The two shaded areas surrounded by the defect are caused by the coil whose heat is captured by the infrared camera. The current, transformed by the thermal image, in the two shaded areas is higher, so the resulting impedance is smaller. The disadvantage is that there is a lot of noise in the low current area. Since the eddy current intensity is mainly concentrated in the vertical area of the defect and gradually decreases to the left and right

sides, the eddy current density on both sides is small and the error is large. In the thermal image processing, the eddy current precision in the low current area is low because only the high current area is concerned, but not the low current area. However, in the impedance distribution, the impedance error obtained in the low current area is large.

The solution is to extract the high current area during the thermal image processing. The impedance distribution is obtained from the high current area, as shown in Fig. 10. The defect profile in the impedance image is clearer and the noise is reduced. The thermal image can only recognize the two high temperature areas of the defect, while the impedance image can recognize the defect profile as shown in Fig. 11.

V. CONCLUSION

In this paper, a defects recognition method based on impedance is proposed. Impedance distribution image can recognize the location and profile of defects more clearly. It is a defects recognition method that combines electrical, magnetic and thermal characteristics of conductive materials.

The traditional method is to extract defect features from the thermal images acquired by infrared cameras. The defect features is the high temperature areas formed by eddy current. Due to the edge effect of the eddy current, when the eddy current encounters a defect, it will bypass the defect and form high temperature areas at the left and right ends of the defect. However, the high temperature areas only exists on the left and right sides of the defect, so it cannot recognize the profile of the defect. The eddy current distribution can be obtained by the thermal images through (1). It clearly describes the flow direction of the eddy current when it encounters a defect. The surface of the conductor sample can be equivalent to the parallel connection of each row of resistances. The electric field intensity of each row is assumed to be evenly distributed. The impedance of each row can be obtained from the electric field intensity and the eddy current density. The impedance distribution can show the defect profile.

Future work will focus on how to eliminate the noise in the low current area. The more accurate impedance distribution will be obtained through optimization algorithm. On the other hand, the electrical impedance tomography (EIT), a medical imaging technique, can be used in defects recognition.

REFERENCES

[1] Avdelidis, N. P. , B. C. Hawtin , and D. P. Almond . "Transient thermography in the assessment of defects of aircraft composites." Ndt & E International 36.6(2003):433-439.

[2] Meola, Carosena . "Nondestructive Evaluation of Materials With Rear Heating Lock-In Thermography." IEEE Sensors Journal 7.10(2007):1388-1389.

[3] Huth, S. , et al. "Localization of gate oxide integrity defects in silicon metal-oxide-semiconductor structures with lock-in IR thermography." Journal of Applied Physics 88.7(2000):4000-4003.

[4]   Ghali, Venkata Subbarao , N. Jonnalagadda , and R. Mulaveesala . "Three-Dimensional Pulse Compression for Infrared Nondestructive Testing." IEEE Sensors Journal 9.7(2009):832-833.

[5]   Ben Weekes and Darryl P. Almond and Peter Cawley and Tim Barden. "Eddy-current induced thermography—probability of detection study of small fatigue cracks in steel, titanium and nickel-based superalloy." NDT & E International (2012).

[6]   Wilson, John , et al. "PEC thermography for imaging multiple cracks from rolling contact fatigue." Ndt & E International 44.6(2011):505-512.

[7]   He, Yunze , M. Pan , and F. Luo . "Defect characterisation based on heat diffusion using induction thermography testing." Review of entific Instruments 83.10(2012):433-439.

[8]   Danny. "Detection and reconstruction of complex structural cracking patterns with electrical imaging." NDT & E International: Independent Nondestructive Testing and Evaluation (2018).

[9]   D. Liu, D. Gu, D. Smyl, J. Deng and J. Du. "Shape Reconstruction Using Boolean Operations in Electrical Impedance Tomography," in IEEE Transactions on Medical Imaging 39.9(2020): 2954-2964.

[10]   Cultrera, Alessandro , et al. "Mapping the conductivity of graphene with Electrical Resistance Tomography." Scientific Reports 9.1(2019):10655.

[11]   Bai, L. , et al. "Spatial and Time Patterns Extraction of Eddy Current Pulsed Thermography Using Blind Source Separation." IEEE Sensors Journal 13.6(2013):2094-2101.

[12]   Zhu, Peipei , et al. "Local Sparseness and Image Fusion for Defect Inspection in Eddy Current Pulsed Thermography." IEEE sensors journal 19.4(2019):1471-1477.

[13]   Cheng, Yuhua , et al. "Stainless Steel Weld Defect Detection Using Pulsed Inductive Thermography." IEEE Transactions on Applied Superconductivity 26.7(2016):1-4.

Telemetry Data-based Spacecraft Anomaly Detection Using Generative Adversarial Networks

Yue Song
School of Automation Science and Electrical Engineering Department
Beihang University
Beijing, China
18810930523@buaa.edu.cn

Jinsong Yu*
School of Automation Science and Electrical Engineering Department
Beihang University
Beijing, China
yujs@buaa.edu.cn

Diyin Tang
School of Automation Science and Electrical Engineering Department
Beihang University
Beijing, China
tangdiyin@buaa.edu.cn

Danyang Han
School of Instrumentation Science and Opto-electronics Engineering
Beihang University
Beijing, China
hdy_daniel@buaa.edu.cn

Sen Wang
National Laboratory of Science and Technology on Aerospace Intelligent Control
Beijing Aerospace Automatic Control Institute
Beijing, China
wangs.idea@foxmail.com

Abstract—The telemetry data of spacecraft is an ultra-high dimensional time series used to indicate on-orbit operation status, and anomaly detection can effectively ensure safety and reliability. Aiming at the characteristics and complex correlation of high dimensional telemetry data, this paper proposes a novel anomaly detection method based on Generative Adversarial Networks (GAN) for telemetry data anomaly detection. Instead of treating each variable independently, our proposed method captures the latent representation amongst multi-dimensional time series. For normal data, the GAN-based anomaly detection method can obtain a reconstructed time series similar to the original time series, learning the probability distribution model of normal data. For abnormal data, the reconstructed time series deviates greatly from the original time series. In the GAN framework, we use Long Short-Term Memory (LSTM) as the network structure of generator for time series reconstruction and discriminator for calculating the probability of being the real time series, which can learn the temporal features of telemetry data. We also propose a novel anomaly score called *GDScore*, which comprehensively considers the reconstruction error of the generator and the output of the discriminator. We conduct experiments with two telemetry datasets, which verifies that our proposed GAN-based anomaly detection method can effectively detect outliers.

Keywords—anomaly detection, telemetry data, multivariable time series, Generative Adversarial Networks, anomaly score

I. INTRODUCTION

When the spacecraft is in orbit, its internal monitoring system obtains sensor information and transmits it to the ground through the telemetry system. Then the telemetry data is sent to the ground operators for monitoring the current operating status of the spacecraft. The telemetry data can be regarded as a high-dimensional time series composed of multiple variables such as temperature and pressure, and the outliers can reflect the failure of acquisition equipment, the damage of transmission system, the degradation of the mechanical and electronic systems, etc. By effectively detecting the outliers in the telemetry data, troubleshooting, remote repair, maintenance and optimization of damaged equipment can be performed [1]. Anomaly detection is of great significance for enhancing the safety and reliability of the spacecraft in all stages of design, production, operation and maintenance.

Currently, expert systems are usually used for anomaly detection of spacecraft telemetry data [2]. The expert system method first performs artificial threshold judgment on the univariable time series, and then combines the correlation between variables to detect the anomaly of the spacecraft system or subsystems [3]. However, with the complexity and diversification of spacecraft types, the telemetry data to be monitored is large and high-dimensional, and the correlation between variables is complex, which leads to the setting of thresholds and correlations through expert experience is inefficient and difficult to achieve. Besides, labeling large amounts of data is difficult and the actual telemetry data generally lacks prior knowledge. Therefore, it is necessary to perform unsupervised multivariate anomaly detection based on the data mining method.

The data mining method can learn the system behavior models for anomaly detection from historical data automatically, rather than from human experts. The data mining-based anomaly detection mainly includes distance-based [4], subspace learning-based [5] clustering-based and deep neural network-based [7, 8] methods. In this paper, we study the deep neural network-based method, which can avoid complex feature engineering and directly extract features from the training data.

Previous studies have shown that in image anomaly detection, such as tumor detection [9], object detection [10], the Generative Adversarial Network (GAN) is effective to capture the probability distribution of normal data through stochastic variables. Inspired by these studies, we first use deep neural network based on GAN to identify outliers in spacecraft telemetry data. The core idea of GAN-based anomaly detection is to learn the latent representation of normal data, which captures normal patterns of multivariate time series.
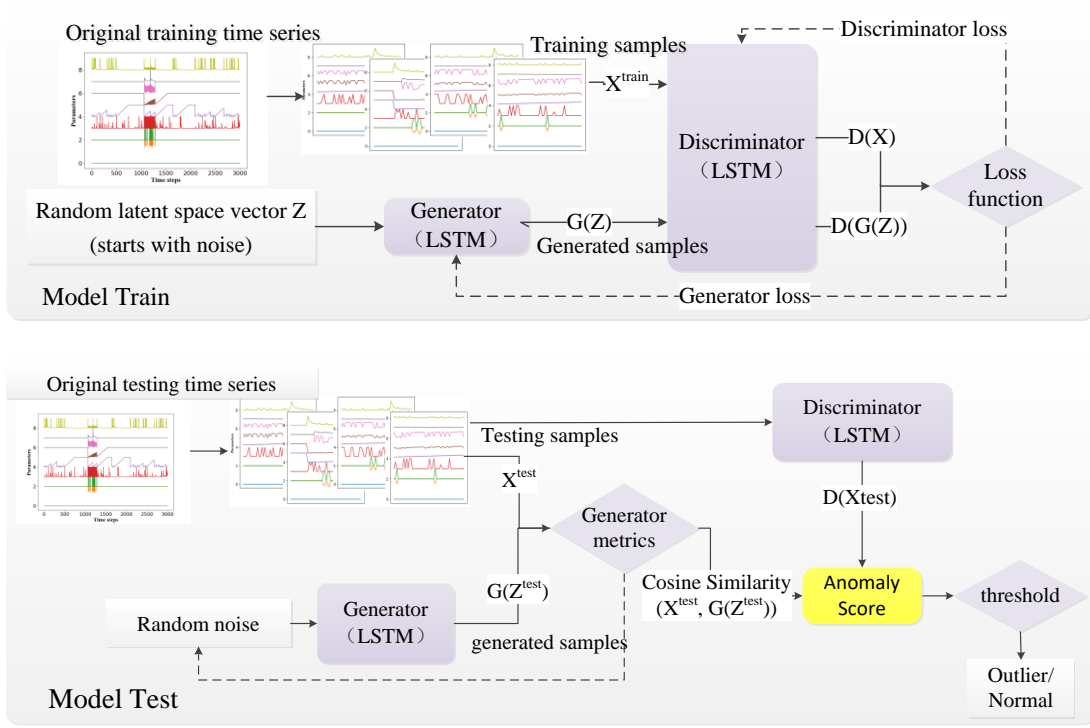
Fig. 1. The overall architecture of GAN

Since telemetry data is a multidimensional time series, the biggest challenge of GAN for telemetry data anomaly detection is how to learn a robust latent representation while considering temporal dependence.

In this paper, we propose a novel telemetry data anomaly detection method using GAN based on Long Short-Term Memory (LSTM) network. Firstly, we split the multidimensional time series of telemetry data into subsequences by sliding window, which enriches the raw time series. Secondly, for the GAN generator and discriminator, we designed the LSTM network structure separately, which can extract the temporal features of the subsequences and predict the state at the next timestamp. Thirdly, we use GAN to reconstruct the subsequences. Finally, We comprehensively consider the reconstruction error of the generator and the output of the discriminator as the anomaly score (called GDScore), which is used to represent the anomaly probability of the next timestamp of current subsequence. After judging the anomaly score by threshold, we can obtain whether each timestamp is an outlier.

The rest of the paper is organized as follows. Section 2 describes the framework overview of GAN-based multivariable anomaly detection. Section 3 specifics on each module in the GAN framework. Experimental results are reported in Section 4. Finally, Section 5 concludes the paper.

II. THE FRAMEWORK OVERVIEW OF GAN

For telemetry data anomaly detection, the objective is to find the outliers in a multivariate time series $\chi \subset R^{T \times N}$, where T is the length of timestamps and N is the number of variates[11]. In order to predict whether the current observation value x_t is outlier, the historical observation values of the previous s_w timestamps $x_{t-s_w:\ t-1}$ are used to predict the anomaly state of x_t. So the original time series is split into a series of sub-sequences by sliding window. Then GAN is used for multivariate time series anomaly detection.

The overall architecture of GAN-based multivariable anomaly detection is illustrated in Fig. 1.

In the model training stage, the training samples are all normal data and the GAN model is used to extract features such as temporal dependence of normal data. Given a random vector in the latent space Z_{train}^k, the generator (G) implicitly defines a generated distribution P_g to fit the original sample distribution P_{data} by mapping Z_{train}^k to the generated samples $G(Z_{train}^k)$. Then $G(Z_{train}^k)$ is input to the discriminator (D) together with the original training samples X_{train}, and D discriminates whether the current input comes from the original sample or the generated sample. The network structure in generator and discriminator are both based on LSTM, which guarantees the ability to effectively extract temporal features and process time series. GAN trains G and D alternately with each other. D is trained to distinguish the original samples X_{train} and the generated samples $G(Z_{train}^k)$ as much as possible, while G is trained to generate samples that are very similar to the original normal samples to mislead D. The entire optimization process can be regarded as a minimax game problem. The loss function of CRGAN is[12]:

$$\min_G \max_D V(D,G) = E_{x \sim p_{data}}[log\, D\,(X_{train})] + E_{x \sim p_g}\left[log\left(1 - D\left(G(Z_{train}^k)\right)\right)\right] \quad (1)$$

where $x \sim p_{data}$ represents the samples from the real data distribution while $x \sim p_g$ denotes the samples from the generated data distribution. $E_{x \sim p_{data}}[log\, D\,(X_{train})]$ represents the expectation that the original samples are judged to real by D while $E_{x \sim p_g}\left[log\left(1 - D\left(G(Z_{train}^k)\right)\right)\right]$ denotes the expectation that the generated samples are judged to fake by D.

The performance of D and G is continuously improved through the adversarial learning. It can be proved that the generated samples by G perfectly fit the original sample

distribution when $P_g = P_{data}$. Since the training samples are all normal sub-sequences, G and D capture the implicit multivariate distribution of normal time series.

In the model testing stage, random noise Z_{test}^k and test samples X_{test} are input into the model to obtain the generated samples $G(Z_{test}^k)$ from generator and the outputs $D(X_{test})$ and $D\left(G(Z_{test}^k)\right)$ from discriminator. Then the anomaly score *GDScore* are calculated based on the reconstruction error of G and the output of D. Finally, the time points in the test dataset are labeled (1 for normal and 0 for anomalous) by judging the threshold τ.

III. THE DETAILS OF GAN-BASED ANOMALY DETECTION

A. Data Pre-processing

Firstly, the multivariate time series $\chi \subset R^{T \times N}$ is split into a train dataset $\chi_{train} \subseteq R^{T_1 \times N}$ and a test dataset $\chi_{test} \subseteq R^{T_2 \times N}$. Note that all the time points in the train dataset must be normal.

Then, the train dataset $\chi_{train} \subseteq R^{T_1 \times N}$ is segmented into a series of subsequences $X_{train} = \{x_i, i = 1,2,\cdots,m\} \subseteq R^{S_w \times N}$ by a sliding window.

Given a window with size s_w and step length s_s, the number of subsequences can be calculated by $m = \frac{T_1 - s_w}{s_s} + 1$. When $T_1 - s_w$ is not divisible by s_s, the last subsequence x_m consists of the last s_w points of the train dataset.

Similarly, the test dataset $\chi_{test} \subseteq R^{T_2 \times N}$ is divided into subsequences $X_{test} = \{x_j, j = 1,2,\cdots,n\}$ by sliding window, where $n = \frac{T_2 - s_w}{s_s} + 1$. To valid the anomaly detection ability of GAN, every point in the test dataset is labeled by binary digits (1 for normal and 0 for anomalous).

B. The Architecture of Generator and Discriminator

After the sliding window processing, a subsequence can be described as H_i, which is a $R^{S_w \times N}$ matrix. To capture the temporal features in a matrix H_i, we use LSTM as the network structure of G and D.

In the generator, the input is a random one-dimensional vector Z of latent space. And then Z is concatenated to a fully connected layer and reshaped into a two-dimensional matrix H_i'. After two LSTM layers and one fully connected layer, H_i' is reshaped into the reconstructed two-dimensional matrix H_i^g, which is the generated time series. The specific network structure of generator is shown in Fig. 2.
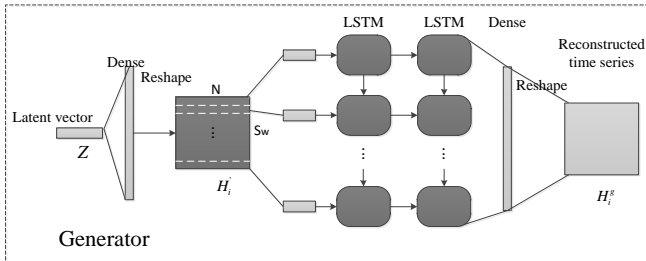
Fig. 2. The network structure of generator

In the discriminator, the input is the real time series H_i or the reconstructed time series H_i^g. After two LSTM layers and one dense layer, the time series is converted to a number. Then the number is calculated by the sigmoid activation function,

and the probability of whether the input time series is a real time series is obtained. The network structure of discriminator is illustrated in Fig. 3.
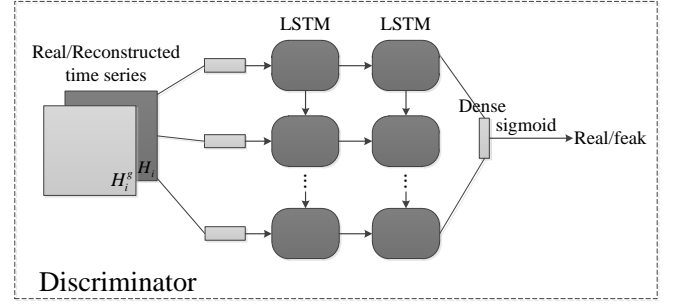
Fig. 3. The network structure of discriminator

C. The Anomaly Score—GDScore

As aforementioned, GAN obtained by training dataset can extract the features of normal time series. For testing dataset, the sub-sequences used to predict normal points are input into G and D to obtain anomaly scores similar to the training dataset, while the scores obtained by sub-sequences used to predict outliers are very different. Therefore, we can comprehensively judge whether the time point is an outlier by measuring the metric after inputting it into G and D. The specific metrics of G and D are as follows:

Generator metric G_m can be calculated by the cosine similarity between the original sample and the generated sample. The lower the cosine similarity, the more abnormal the original sample, then the more abnormal the time point predicted by the subsequence. G_m can be calculated by:

$$G_m = Cos\,i\,ne\ similarity \left(X_{test}, G(Z_{test}^k)\right)$$
$$= \frac{\sum_{i=1}^{n} x_i \times G(Z_{test}^k)_i}{\sqrt{\sum_{i=1}^{n}(x_i)^2} \times \sqrt{\sum_{i=1}^{n}[G(Z_{test}^k)_i]^2}} \quad (2)$$

The discriminator metric D_m can be measured using the predicted value output by the sigmoid function of the last layer, which directly represents the probability of being normal / outlier at this time point. D_m can be described as:

$$D_m = D(X_{test}) \quad (3)$$

In order to synthesize the results of generator metric and discriminator metric, this paper proposes a new calculation method of anomaly score (named *GDScore*). *GDScore* can be described as follow:

$$GDScore = 1 - \frac{G_m + D_m}{2} \quad (4)$$

IV. EXPERIMENTS

We conduct the experiments of GAN-based anomaly detection on two public telemetry datasets.

A. Datasets and Evaluation Metrics

SMAP (Soil Moisture Active Passive satellite) and MSL (Mars Science Laboratory rover) are two public telemetry datasets from NASA. SMAP contains a time series of 25-dimensional variables, and the length of all the time series of the 55 channels is 429735. The MSL contains a time series of 55-dimensional variables, and the total length of all the time series of the 27 channels is 66709 [13]. Table I shows the details of two datasets.

TABLE I.   THE DETAILS OF DATASETS

Dataset name	Number of channels	The total length of time series	Number of variables
SMAP	55	429735	25
MSL	27[a]	66709	55

SMAP and MSL datasets are divided into a training set and a test set, and the test set is labeled with 0 or 1. It should be noted that the training set can only contain normal data, and the test set contains normal and anomalies data. Then we set window size $s_w = 30$ and step length $s_s = 10$ to split the original time series into subsequences.

We use Precision, Recall and F1-score to evaluate the performance of the GAN-based anomaly detection method, where $Precision = \frac{TP}{TP+FP}$, $Recall = \frac{TP}{TP+FN}$ and $F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$ [14].

B. Experimental settings and platform

We set the length of the input latent vector Z of the generator to 100. The two-layer LSTM of the generator and discriminator respectively contain 80 and 40 cells, and dropout is added after each layer of LSTM. Other parameter settings are shown in the Table II.

TABLE II.   THE OTHER SETTINGS OF GAN

Generator optimizer	Discriminator optimizer	Loss function	Batch size	Epochs
Adam	Adam	Cross entropy	64	50

Because the performance of the generator is weaker than the discriminator during training and the discriminator converges faster, so we set the generator's initial learning rate to be greater than the discriminator, which is 0.002 and 0.0001 respectively.

The GAN-based anomaly detection method are implemented in Python 3.7 with the help of deep learning library TensorFlow-GPU 1.8.0. Experiments are performed on a Huawei G5500 Linux server with NVIDIATesla V100 GPU.

C. Results

To demonstrate the effectiveness of the GAN-based telemetry data anomaly detection algorithm (LSTM-GAN), we compared it with two other unsupervised anomaly detection methods. They are isolated forest (Iforest) and LSTM-based Variational Autoencoder (LSTM-VAE). Iforest divides trees to achieve the purpose of clustering normal data and abnormal data [6]. LSTM-VAE is another method to represent normal data by using stochastic variables, and it includes two parts, an encoder and a decoder. Both the encoder and the decoder use the LSTM structure [8].

After conducting network training and testing on GAN, we obtained the results of anomaly score *GDScore* of the test dataset. When the threshold τ is set to optimal, the Precision, Recall and F1-score are calculated on the SMAP and MSL datasets for the three methods, which are shown in the Table III.

Depicted in the Table III, The F1 obtained by LSTM-GAN are higher than LSTM-VAE and Iforest on both SMAP and MSL datasets, which proves that LSTM-GAN outperforms the other methods on the multivariable time series anomaly

detection. Also, the difference between Precision and Recall of LSTM-GAN is not large, indicating that LSTM-GAN can balance the false alarm rate and the missed detection rate, so that the anomaly detection results of telemetry data can reach the comprehensive optimal.

TABLE III.   THE PERFORMANCE OF THE THREE METHODS

Methods	SMAP(%)			MSL(%)		
	Precision	*Recall*	*F1*	*Precision*	*Recall*	*F1*
Iforest	86.75	89.59	88.14	89.73	90.26	89.99
LSTM-VAE	74.16	97.76	84.34	88.67	91.17	89.89
LSTM-GAN	88.95	92.19	90.54	91.61	92.71	92.16

In addition, because the Iforest method ignores the time dependence between data, it is only suitable for numerical anomalies, but not for time-dependent anomalies. Comprehensively speaking, the LSTM-GAN method performs well in both effectiveness and applicability.

Then, we consider the influence of the threshold on the anomaly detection results, and find the best threshold by searching the [0, 1] interval. Fig. 3 and Fig. 4 respectively show the effect of threshold on the anomaly detection results of SMAP and MSL data sets.
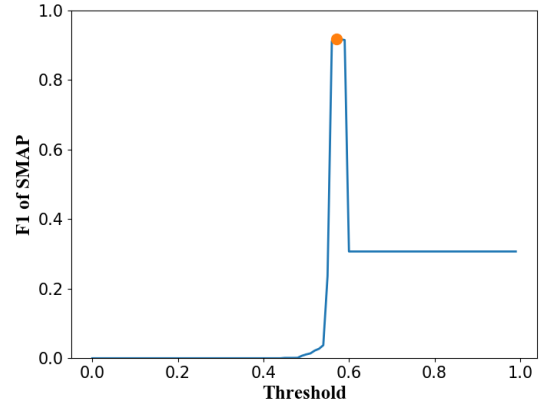
Fig. 4. The effect of threshold on MSL dataset
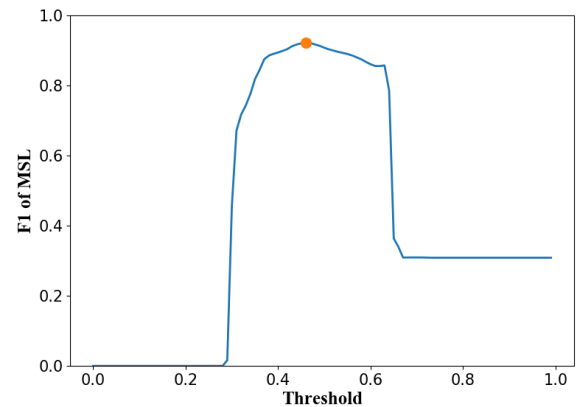
Fig. 5. The effect of threshold on MSL dataset

When the threshold is too small, Recall will be small, resulting in F1 being too small; when the threshold is too large, Precision will be small, causing F1 to be too small. So, only when the threshold is set moderately, can F1 be guaranteed to be optimal. As the threshold increases, F1 converges, and the final convergence value is related to the proportion of outliers in the dataset. As shown in Fig 3 and Fig. 4, when the

threshold for SMAP is 0.58, the optimal F1 is 90.54%; when the threshold for MSL is 0.46, the optimal F1 is 92.16%.

V. CONCLUSIONS

This paper proposes a novel telemetry data anomaly detection method based on GAN, using the LSTM as the structure of generator and discriminator. This method makes three significant contributions. Firstly, GAN-based telemetry data anomaly detection can automatically extract time features and correlation features of multiple variables, which solves the problem of lack of labels in telemetry data. Secondly, LSTM structure is introduced to learn a robust latent representation of telemetry data while considering temporal dependence. Finally, a new anomaly score called *GDScore* is proposed to consider the reconstruction error of generator and the output of discriminator simultaneously. The experimental results show that GAN-based anomaly detection method preforms well on the telemetry datasets SMAP and MSL compared with other methods.

ACKNOWLEDGMENT

REFERENCES

[1] T. Yairi, Y. Kawahara, R. Fujimaki, Y. Sato and K. Machida, "Telemetry-mining: a machine learning approach to anomaly detection and fault diagnosis for space systems," *2nd IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT'06)*, Pasadena, CA, 2006, pp. 8 pp.-476.

[2] C. Chang, W.Nallo, R.Rastogi, D.Beugless, F.Mickey, and A.Shoop. Satellite diagnostic system: An expert system for intelsat satellite operations. In Proc. IVth European Aerospace Conference (EAC), pages 321–327, 1992.

[3] M. Rolincik, M. Lauriente, H. C. Koons, and D. Gorney. An expert system for diagnosing environmentally induced spacecraft anomalies.

[4] E. Khalastchi, M. Kalech, G. A. Kaminka, and R. Lin, "Online data driven anomaly detection in autonomous robots," Knowl. Inf. Syst., vol. 43, no. 3, pp. 657–688, Jun. 2015.

[5] Y. He, Y. Peng, S. Wang and D. Liu, "ADMOST: UAV Flight Data Anomaly Detection and Mitigation via Online Subspace Tracking," in *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 4, pp. 1035-1044, April 2019.

[6] Liu, Fei Tony , K. M. Ting , and Z. H. Zhou . "Isolation-Based Anomaly Detection." Acm Transactions on Knowledge Discovery from Data 6.1(2012):1-39.

[7] Kieu, Tung, Bin Yang, and Christian S. Jensen. "Outlier detection for multidimensional time series using deep neural networks." 2018 19th IEEE International Conference on Mobile Data Management (MDM) . IEEE, 2018.

[8] Su, Ya, et al. "Robust Anomaly Detection for Multivariate Time Series through Stochastic Recurrent Neural Network." Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining . 2019.

[9] D, Yanan Ruan A , et al. "MB-FSGAN: Joint segmentation and quantification of kidney tumor on CT by the multi-branch feature sharing generative adversarial network." Medical Image Analysis (2020).

[10] A, Robert Skilton , and Y. G. B . "Combining object detection with generative adversarial networks for in-component anomaly detection." Fusion Engineering and Design 159.

[11] Li, Dan, et al. "Mad-gan: Multivariate anomaly detection for time series data with generative adversarial networks." International Conference on Artificial Neural Networks . Springer, Cham, 2019.

[12] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In NIPS'2014.

[13] Kyle Hundman, Valentino Constantinou, Christopher Laporte, Ian Colwell, and Tom Soderstrom. 2018. Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18). ACM, New York, NY, USA, 387–395.

[14] C. Sammut and G. I. Webb, Eds., Encyclopedia of Machine Learning and Data Mining. Springer, 2017.

In Proc. of 5th Annual Space Operations and Applications Research Symposium, pages 36–44, 1992.

Workpiece Detection Based on Image Processing and Convolutional Neural Network

Wocheng Chen
School of Mechanical and Automotive Engineering
South China University of Technology
Guangzhou, China
2212764765@qq.com

Yi Lei
School of Mechanical and Automotive Engineering
South China University of Technology
Guangzhou, China
645505804@qq.com

Min Liu
School of Mechanical and Automotive Engineering
South China University of Technology
Guangzhou, China
359822279@qq.com

Junming Zhang
School of Mechanical and Automotive Engineering
South China University of Technology
Guangzhou, China
297779460@qq.com

Xifan Yao*
School of Mechanical and Automotive Engineering
South China University of Technology
Guangzhou, China
mexfyao@scut.edu.cn

Abstract—**This paper presents a method on workpiece detection based on image processing and convolutional neural network(CNN). Firstly, four extreme points and center point of the workpiece are detected by image processing technologies such as canny edge detection operator, morphological processing and denoising processing. And the predicted boxes fitting the shape of image is generated. Then, according to the coordinates of the extreme points, the image with a single workpiece is cut out, and a novel CNN named workpiece-net(wp-net) is created to classify the object. As a result, the accuracy of image cutting is 0.9986, the average Intersection over Union(IoU) is 0.9235; the parameter size of wp-net is 98.25K and the average precision of classification is 0.9883. In addition, the average recall of classification is 0.9877 and the speed of classification is 0.1243s/fps without Graphics Processing Unit(GPU) and multithreading. Compared with the pure deep learning method, this method can detect more accurate coordinates which are consisted of extreme points. At the same time, the number of wp-net parameters and the complexity of the model structure used for classification are so far less than the popular deep neural network for detection that it can be easily deployed in embedded devices with limited storage space and computing power.**

Keywords—workpiece detection, image processing, convolution neural network, extreme points

I. INTRODUCTION

With the advancement of industry 4.0, more and more personalized customization and small-batch production scenarios appear in the industrial manufacturing process, which makes the detection of workpieces turn from rule and single-target task to multi-target, multi-category and irregular target task [1]. Workpiece detection, including location and classification of workpieces, is mainly realized by machine vision. Traditional detection methods, represented by the method proposed by n. DALAL et al. [8], can effectively detect images that contain only single category, but it is difficult to detect multi-category and multi-target images. This situation has not been improved until the rise of deep learning in recent years. The methods based on deep learning usually show better performance than traditional methods in target detection and image classification, and can effectively achieve multi-category and multi-target detection. So it has been widely used in various fields. For example, compared with the traditional method, emerging neural network models of YOLO's family and R-CNN's family, which are widely used, have a greater improvement in accuracy of detection and classification [2-4]. However, on the one hand, these emerging methods for detection of target have complex network structure and huge parameter scale, which have high requirements for training and deployment equipment. On the other hand, The results may not be accurate enough, especially for some high-precision positioning tasks, such as some stereo vision tasks because their final detection results are the coordinates of the upper left corner and the lower right corner of the rectangular box (as shown in Fig. 1).

Fig. 1. Schematic diagram of typical depth convolution network detection box

In 2019, ExtremeNet was proposed [5]. This method detects four extreme points and center points on the target, including upper part, lower part, left part and right part, and gets octagonal box which is more suitable to the target contour. However, similar to other popular pure deep learning methods, ExtremeNet is complex, and its parameter scale is huge, up to about 200M. Recently, some researchers [6] combined deep learning and image processing technology to detect workpieces. Firstly, the detection boxes of the workpieces are obtained by YOLOv3, and raw images are cropped to get single-workpiece images according to detection boxes. Then, the traditional image processing technology is used to obtain the coordinates of the extreme points of the workpieces in the single-workpiece images. Finally, a more suitable detection box is generated by extreme points. Compared with

EtremeNet, YOLOv3 has fewer parameters and faster detection speed [2]. However, YOLOv3 has about 61.62M amount of parameters and needs a device with high-performance GPU for training and deployment, which is not able to be deployed in many scenarios such as in embedded devices with limited computing power and storage space. In view of the above problems, this paper proposes a novel approach, as shown in Fig. 2.

Firstly, we use a series of image processing methods to detect the extreme points of the workpiece images, and cut out single-workpiece images according to the coordinates of the extreme points. Then, we create a light-weight CNN called workpiece-net (wp-net) to classify the single-images of workpieces. The wp-net consists of 15 layers, whose parameters size is 98.25K. The average precision of classification is 0.9883, and the average recall is 0.9877. Through the combination of traditional image technology and deep learning, we can get high accuracy of location and small scale of the model parameters compared with popular deep learning models.
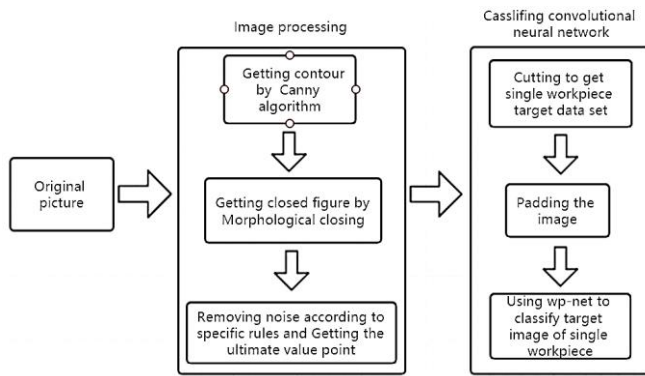
Fig. 2.   Component diagram of workpiece detection

II. RELATED WORK

Workpiece detection includes location and classification. Before the emerging of deep learning, the popular workpiece detection methods include template matching algorithms and learning algorithms. The template matching algorithms obtain the contour of the target by processing the pixels and then matches with the preset template to identify the target workpiece. For example, Huang et al. proposed to use Hu matrix invariance to extract the contour of workpieces and then match with the preset templates, so as to identify the workpiece category [7]. This method can effectively detect the contour and the category of the workpieces, but it is not able to effectively identify the multi-category objects and the objects that are not in the preset templates. In addition, we need to design particular algorithm steps and templates for specific artifacts and scenes. Therefore, this kind of method is usually used in the limited scenes with few kinds of workpieces, high discrimination, fixed objectives and stable production environment. In the field of learning algorithm, the most representative algorithm is the method of HOG combined with SVM proposed by Navneet Dalal et al [8]. This method calculates and statistics the gradient direction histogram of the local area of the image to construct features, and then combines support vector machine(SVM) classifier to realize pedestrian detection. In addition, this method can be used for image classification. However, This method is usually used to detect single-category images because it is difficult to achieve multi-category and multi-target detection.

In addition, learning algorithms often need to manually create relevant features for different datasets, such methods are difficult to achieve excellent results in multi-category and multi-objective tasks. For this reason, the research on workpiece task using learning algorithm mainly focuses on image classification task, representing by machine learning such as SVM and its improved algorithm [9,10]. In recent years, the emerging method based on deep learning can automatically learn the relevant features of the images. In addition, it is more robust to noise and has higher accuracy than the traditional machine learning methods. For example, the deep learning algorithms for target detection such as YOLOv3 are able to detect multi-category and multi-target images with high precision and generate the rectangular box of the target. The ExtremeNet, which was proposed in 2019, can directly detect four extreme points on the targets, which improve the accuracy of location. However, the network structure is complex and its amount of parameters is about 200M. Because the training and deployment require very high computing power, the deep CNN is difficult to be applied in many scenarios. Reducing the complexity of the models for detection is an important trend of current research. At present, there are methods such as model pruning and network decomposition to reduce the complexity of the model [11]. For example, Li et al. put forward [12] the method of tailoring filters with sum of filter weight as significance measure, which can reduce the amount of model parameters and the complexity of the model to a certain extent, but the effect of the method is still limited. In addition, tailoring model often leads to performance degradation. Another way to reduce the complexity of deep convolution neural network is combining image processing technology with deep learning. Recently, some researchers proposed a method that combines deep learning with image processing technology [6], which get the detection boxes by YOLOv3 and crop raw images to get single-workpiece images according to detection boxes. Then, the traditional image processing technology are used to obtain the coordinates of the extreme points of the workpieces in the single-workpiece images. Finally, a more suitable detection box is generated by extreme points. As a result, the model has about 61.62M amount of parameters and the network complexity is greatly reduced compared with ExtremeNet that has about 200M amount of parameters, but it is still at a high level. Obviously, image detection that includes location and classification, is more complex than image classification. So the deep learning model used for image detection is far more complex than that for classification. And the models for image detection have much larger amount of parameters. Based on the combination of image processing technology and CNN, we use the former to detect the extreme points in the positioning task and uses the latter to complete the classification of workpieces respectively. Compared with the deep CNN model for detection, the model for classification is so simple that it can be deployed in embedded devices or mobile devices with limited computing power.

III. WORKPIECE DETECTION METHOD

As mentioned above, the workpiece detection needs to realize workpieces location and classification. This study proposes a novel method which combines a series of image processing technologies with deep learning. Through simple and fast technologies of image processing, workpiece location is completed. And light-weight model is used to accurately classify multi-category tasks, and obtain more accurate coordinates composed of extreme points and higher

classification accuracy with a simple model structure of few parameters. There are two steps to achieve the workpiece detection: using a series of image processing methods to achieve the target location, and using the novel CNN called wp-net to achieve the target.

A. Workpiece Location

Firstly, our method uses canny algorithm to get the contour. After obtaining the contour of the workpiece image, in order to remove the noise, we use morphological filtering to fill the contour gap to make obviously difference between the contour of workpiece and noise. Then, this method extracts the coordinates of extreme points on the contour and calculates the center points. As shown in Fig. 3, the processed image has some noise that is small and discontinuous contour.
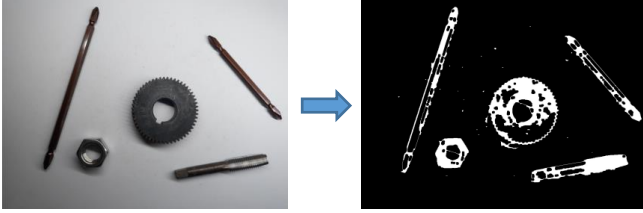
Fig. 3.   Using Canny algorithm to process workpiece contour

According to the image analysis after contour extraction, the noise is discontinuous and fragmented. Therefore, by calculating the number of contour pixels and contour areas, such a contour with the number of pixels and the contour area less than the preset thresholds is taken as noise, which is able to remove most of the noise.

However, in the actual processing, there are some exceptional cases that cause positioning errors. As shown in Fig. 4, for the image with inner holes, the contour of the inner hole is detected and causes errors.
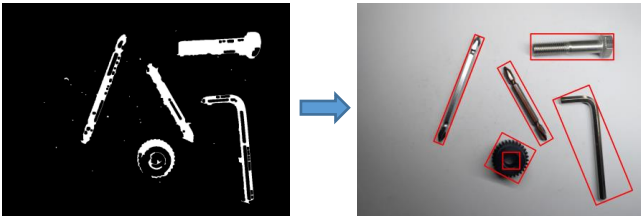
Fig. 4.   Inner hole of workpieces detected

For dealing with this problem, the following rule is set: when the distance between the center points of two predicted boxes is less than the threshold and the coordinates of four extreme points of a box are completely covered by another box, the predicted box with a smaller area is treated as noise. Through this rule, such a problem can be solved well. Finally, by cutting raw images with the extreme points, we get the initial single-workpiece classification dataset.

B. Workpiece Classification

In consideration of the excellent performance of CNN in image classification, this study uses CNN to achieve the task of workpiece classification. After obtaining the initial classification dataset, we need to adjust images to a fixed size because that the common classification neural network accepts fixed size input. The popular methods are resizing and padding. Considering that resizing may lead to image deformation, this study uses padding and finally obtains the image with the size of 512 * 512. Then, According to the

complexity of the job for classification and the scale of the dataset, the so called wp-net with suitable depth is designed. As shown in Fig. 5, the network is composed of 15 layers and accepts images with the size of 512 * 512 as inputs. The model has 98.25K amount of parameters.

a. The left side is the top part and the right side is the bottom part.

Fig. 5.   Wp-net structure

The wp-net includes typical CNN layers such as convolution, pooling and activation function [13]. In addition, the last two layers of the network are designed as full connection layer to achieve classification [14] and each neuron in the last layer of full connection layer outputs the probability of each workpiece category. Besides, the batch-normalization (BN) layer [15] is added to the structure to improve the generalization ability of the network. Wp-net uses Adam [16] as the optimization algorithm and uses the linear rectification activation function to prevent the gradient from disappearing and improve the training speed [17]. In addition, considering the network output is probability, cross entropy loss function [18] is used. After the completion of network construction, the wp-net is trained and tested directly by using the dataset of workpieces. The training process is carried out on the CPU of i5 processor without GPU and multiprocessing. The batchsize is 16 and epochs is 5.

IV. EMPIRICAL ANALYSIS

As stated above, this study firstly uses a series of image processing technologies to detect the extreme points of the workpiece images and cut the pictures of workpiece according to the extreme points. And the single-workpiece dataset is produced referring to the format of MNIST [19], which is used to train the wp-net. Then, we use the trained wp-net to examine the test set. Finally, the above two parts are integrated to realize the location and classification of workpiece images.

A. Extreme Points Detection and Image Cutting Experiments

In this study, a dataset of workpieces that has ten categories is established, including bolts, nuts, screwdrivers, screw dies, drills, wrenches, l-wrenches bearings, taps and gears. This dataset has 10488 images, each of which contains at most seven categories and nine workpieces. Based on

OpenCV, the whole process of extreme points detection and image cutting is completed. After removing the noise, we collect 38750 single-workpiece images by cutting raw images and the accuracy of cutting is 0.9986. In addition, we use labelImg to label the original images and get labels in the format of COCO [20]. Finally, the Intersection over Union(IoU) [21] is introduced to reflect the accuracy of the location. As shown in the following formula, the IoU reflects the accuracy of predicted boxes location by calculating the ratio between the *pre_ box_ Area* and *true_ box_ Area*, where the former is the area of predicted box, and the latter is the area of real box.

$$IoU = \frac{pre\_box\_area \cap true\_box\_area}{pre\_box\_area \cup true\_box\_area} \quad (1)$$

As a result, the average IoU between all predicted boxes and real boxes is 0.9235. Under the condition of i5 processor without multiprocessing, the cutting speed of images is 0.0766 s/fps.

B. Using wp-net to Classify Single-Workpiece Images

After cutting the raw images to get single-workpiece images, each image is padded. As a result, we get the fixed size images, which is divided into the 10 categories. And we get the dataset in the format of MNIST. Then, 27125 images are randomly selected as the training set, 7750 images as the validation set and 3875 images as the test set. Next, wp-net is trained five epochs based on the training set by PyTorch. As shown in Fig. 6, the average classification accuracy of all categories of the verification set is 0.9881 at the fifth epoch. The average classification accuracy on the test set is 0.9880.

In addition, in order to fully consider the difference of different categories in the multi-classification problem, the precision and recall are introduced. As shown in the Table I, the precision and recall of each category are at a high level. The average precision and average recall of all categories are 0.9883 and 0.9877 respectively. It is found that the recall of screwdriver is 0.9441.In addition, the precision and recall of the other parts are above 0.96.
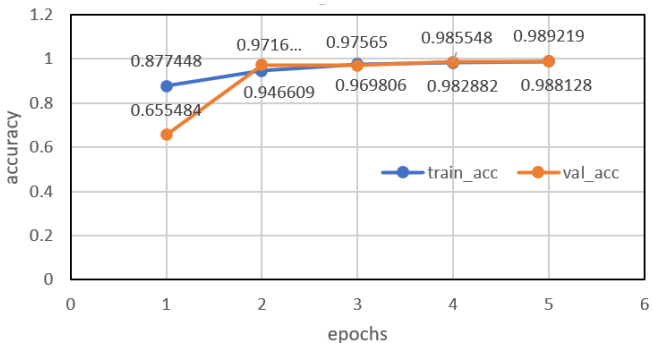
Fig. 6. Accuracy of raining and verification sets with epochs

It is speculated the reason is that the screwdriver has a long strip shape and its image area is large, which often contains part of the other categories that cause misclassification.

As for the complexity and classification speed, the wp-net has 98.25K amount of parameters, which is far less than the popular deep learning network model for detection. As a result, the model is easy to be deployed in embedded devices with limited computing power and storage space. In addition, under the condition of i5 processing and without GPU, the image classification speed is 0.1243 s/fps.

TABLE I. AVERAGE PRECISION AND RECALL OF EACH CATEGORY

Workpiece category	Average precision	Average recall
L-wrench	0.9808	0.9903
wrench	1.0000	0.9918
Screw die	1.0000	1.0000
gear	0.9973	1.0000
nut	0.9976	0.9953
bolt	0.9894	0.9766
driver	0.9826	0.9441
Tap	0.9716	0.9817
bear	0.9949	1.0000
drill	0.9688	0.9975
all	0.9883	0.9877

C. Integrating Location and Classification

As shown in Fig. 7, after integrating workpiece location and classification, we are able to get the extreme points and categories of multi-category and multi-target workpiece images. And the predicted boxes are more suitable for the target contour, which means that the location results will be more accurate.
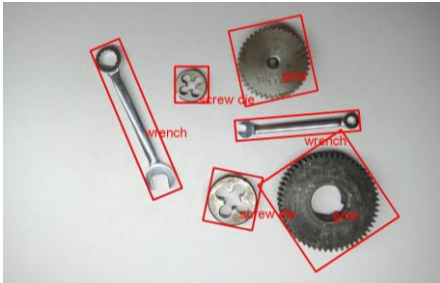
Fig. 7. Demo of the proposed method

Based on the same dataset, we use i5-7200u CPU of Intel to do some comparative study on several methods. In the study, we compared the method that combination of deep learning and image processing [6], the excellent deep learning object detection model named YOLOv3 [2] and the classical image processing method named HOG [8]. The results are shown in Table II. It can be seen that when recall is close to other methods, the model parameters and single frame detection time of our method are far less than other methods. In the task of clamping workpiece for a manipulator, a more accurate clamping strategy can be given according to the coordinates of extreme points and center point.

TABLE II. RESULTS OF DIFFERENT METHOD

Method	Recall	Params (M)	Time (s/fps)	Result of detection	Detection object
YOLOv3	0.9794	61.62	0.6481	Rectangular box	Multiple categories
YOLOv3 with image processing and pruning	0.9891	15.45	0.6675	Extreme points	Multiple categories
HOG+SVM	/	/	/	Rectangular box	Single category
our method	0.9877	0.098	0.2245	Extreme points	Multiple categories

V. CONCLUSION

This study has proposed a workpiece detection method which combines image processing technology with deep

learning. Firstly, we use a series of image processing algorithms to obtain the contour of the workpiece and coordinates of extreme points, which means more accurate location. Then, we use a novel CNN called wp-net to classify the single-workpiece target. Experimental results show that the image processing technology can get high IoU for workpiece location. At the same time, the complexity of model and amount of parameters are greatly reduced compared with pure deep CNN for detection.

REFERENCES

[1] X. Yao, J. Zhou, J. Zhang, and C. R. Boer, "From Intelligent Manufacturing to Smart Manufacturing for Industry 4.0 Driven by Next Generation Artificial Intelligence and Further On." 2017 5th International Conference on Enterprise Systems (ES), 2017, pp. 311–318.

[2] J. Redmon, A. Farhadi, "YOLOv3: An Incremental Improvement." ArXiv Preprint ArXiv:1804.02767, 2018.

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, 2017, pp. 1137–1149.

[4] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection." ArXiv Preprint ArXiv:2004.10934, 2020.

[5] X. Zhou, J. Zhuo, Krähenbühl, and Philipp, "Bottom-Up Object Detection by Grouping Extreme and Center Points." 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 850–859.

[6] Y. Lei, X. Yao, W. Chen, J. Zhang, J. Mehnen, and E. Yang, "Multiple Object Detection of Workpieces Based on Fusion of Deep Learning and Image Processing." IEEE World Congress on Computational Intelligence 2020, 2020. in press.

[7] C. Huang, D. Chen, and X. Tang, "Implementation of Workpiece Recognition and Location Based on Opencv." Computational Intelligence and Design (ISCID), 2015 8th International Symposium On, vol. 2, 2015, pp. 228–232.

[8] N. Dalal, and B. Triggs. "Histograms of Oriented Gradients for Human Detection." 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, 2005, pp. 886–893.

[9] L. Yang, M. Chong, C. Bai, and J. Li, "A Multi-Workpieces Recognition Algorithm Based on Shape-SVM Learning Model." Journal of Physics: Conference Series, vol. 1087, no. 2, 2018, p. 22025.

[10] L. Yang, Y. Dong, J. Zhuang, and J. Li, "A Recognition Algorithm for Workpieces Based on the Machine Learning." 2018 11th International Symposium on Computational Intelligence and Design (ISCID), 2018, pp. 371–375.

[11] J. Rongrong, L. Shaohui, C. Fei, W. Yongjian, and H. Feiyue, "Deep Neural Network Compression and Acceleration: A Review." Journal of Computer Research and Development, vol. 55, no. 9, 2018, p. 1871.

[12] H. Li, A. Kadav, I. Durdanovic and H. Samet, "Pruning Filters for Efficient ConvNets." ICLR 2017: International Conference on Learning Representations 2017, 2017.

[13] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning. 2016.

[14] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks." Communications of The ACM, vol. 60, no. 6, 2017, pp. 84–90.

[15] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." Proceedings of The 32nd International Conference on Machine Learning, 2015, pp. 448–456.

[16] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization." ICLR 2015: International Conference on Learning Representations 2015, 2015.

[17] K. Hara, D. Saito and H. Shouno, "Analysis of Function of Rectified Linear Unit Used in Deep Learning." 2015 International Joint Conference on Neural Networks (IJCNN), 2015, pp. 1–8.

[18] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks." Communications of The ACM, vol. 60, no. 6, 2017, pp. 84–90.

[19] L. Deng, "The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]." IEEE Signal Processing Magazine, vol. 29, no. 6, 2012, pp. 141–142.

[20] T.-Y. Lin, M. Maire, S. J. Belongie, J. Hays, P. Perona, and D. Ramanan, et al. "Microsoft COCO: Common Objects in Context." European Conference on Computer Vision, 2014, pp. 740–755.

[21] A. Rahman and Y. Wang, "Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation." International Symposium on Visual Computing, 2016, pp. 234–244.

A Bayesian Fusion Model for Heart Rate Annotations

Jianan Di
School of Instrument Science and Engineering
Southeast University
Nanjing, China
220183264@seu.edu.cn

Jianqing Li*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
ljq@seu.edu.cn

Chengyu Liu*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
chengyu@seu.edu.cn

Abstract—**Lack of gold standard of truth is a critical problem in medical field. Consensus extraction from noisy labels to estimate ground truth is a common approach. However, significant inter-observer variability and various biases limit the accuracy. For this challenge, we proposed an unsupervised Bayesian framework to aggregate multiple heart rate (HR) annotations from the electrocardiogram (ECG) signal and infer the underlying ground truth and precision of annotators when the ground truth is not available. We further introduced the bias and precision adjustment factor of annotator to model the annotator's performance in each instance. Test on the 2014 PhysioNet/ Computing in Cardiology challenge database showed that the proposed model had an improvement not only over the existing aggregation model (EM-R method), mean and median strategies (the highest of 10.36%，20.26% and 12.43%) but also to the best single annotator (13.75%) used in this paper.**

Keywords—Annotation aggregation, ECG, HR estimation, Bayesian probability, unsupervised learning.

I. INTRODUCTION

Labeling methods are divided into manual and algorithm labeling. Manual labeling is time-consuming and expensive, especially for large-scale data. Algorithm labeling has high efficiency and can work in real-time, which is a key component for ECG devices. However, algorithm is easily affected by signal noise and its accuracy is not as high as manual annotation. Thus, it is necessary to build a model by fusing the labels of multiple algorithm annotators to estimate the underlying true value.

Cardiovascular diseases (CVDs) are the leading cause of death worldwide [1], which has led to a large number of studies on ECG signals [2]. HR monitoring is an important part of assessment and diagnosis of CVDs [3]. At present, the detection algorithm of HR has developed to a relatively mature stage [4]. However, these algorithms are not ideal for poor quality ECG signals [5], which has a great impact on the accuracy of HR monitoring. Therefore, a probabilistic model of fusing HR annotations but no absolutely gold standard was proposed, which can infer the underlying ground HR value and estimate each annotator's ability. Simple methods of forming aggregate labels, such as the mean, median voting and majority voting strategies, generally perform well only if there are a large number of annotators available [6]. Dawid and Skene [7] presented a more effective model early addressing the doctors' diagnostic differences caused by the inconsistent expressions of patients, and firstly used EM algorithm to obtain maximum likelihood estimations of the parameters. Carpenter [8] proposed a complete Bayesian version of the model [7], focusing on binary models of the two categories. Raykar et al. [9] proposed Bayesian posterior estimation of the model and applied it to the classification of benign or malignant breast tissue in CAD diagnosis. And they extended

this model from binary data to continuous data, which is a benchmark method in this field. Experimental results indicated that the model was superior to the majority voting. Later, many researchers made improvements and extensions based on it [9]. Welinder and Perona [10] proposed an online algorithm to estimate the "ground truth" of image properties and the annotators' abilities. The model can handle binary, multi-valued, and continuous-valued annotations, but didn't model the bias of annotators.

To sum up, research of fusion model in the field of ECG annotation is less, and it is vital to obtain high-quality HR annotations for wearable ECG monitoring. In contrast to previous works, this article proposed a Bayesian framework for aggregating multiple HR annotations of ECG, which took into account the precision and bias of the individual annotators. Furthermore, we compared the model with commonly mean, median method and existing fusion model for HR labels [11]. The rest of the paper is organized as follows. Section II describes our IBPP model. Section III shows the applications in the 2014 PhysioNet/Computing in Cardiology challenge (CinC) ECG database. Finally, section IV concludes the paper.

II. METHODS

A. An Improved HR Estimation Model Based on Bayesian Prior Probability (IBPP)

The illustration of the IBPP model presented in this paper is shown as follows.
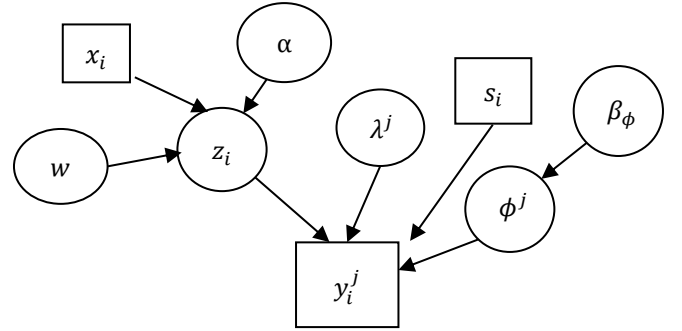
Fig. 1. Graphical representation of the IBPP model: y_i^j corresponds to the annotation provided by the jth annotator for the ith record, and it is modelled by the z_i (the unknown underlying ground truth), the ϕ^j (bias), and the $1/s_i\lambda^j$ (precision with adjustment factor). z_i is drawn from a Gaussian distribution with parameters mean w^Tx_i (x_i corresponds to feature vector for the ith ECG sample, w is regression coefficient) and variance $1/\alpha$ (α is drawn from a Gamma distribution). ϕ^j is modelled from a Gaussian distribution with variance of β_ϕ. β_ϕ and λ^j are drawn from a Gamma distribution respectively. The square box represents the known observations and the circular box represents the unknown variables.

Suppose that there are N signal records and R annotators. y_i^j corresponds to the annotation of the ith record provided by the jth annotator, and z_i represents the unknown underlying ground truth(ground truth means the actual HR value). It is

---

assumed that y_i^j is a noisy version of z_i, which follows a Gaussian distribution with the mean of z_i and the variance of $1/\lambda_j$. Since it is assumed that all the annotator is the independent and identically distributed, on the basis of the central limit theorem, the labels they provide converge to a Gaussian distribution. Furthermore, the annotators' bias ϕ_j are put into consideration, the conditional probability of estimating y_i^j can be written as:

$$P(y_i^j | z_i, \phi_j, \lambda_j) = N(y_i^j | z_i + \phi_j, 1/\lambda_j) \qquad (1)$$

where ϕ_j, λ_j are considered to be constants for the jth annotator, i.e. all annotators are assumed to have consistent but usually different performances throughout records. λ_j represents the precision annotated by the jth annotator. The meaning of ϕ_j is mean error between annotation and underlying ground truth. Besides fixed features of annotators, the distribution of biases might be influenced by datasets. In the absence of any knowledge of the underlying distribution of biases, ϕ_j is assumed to be drawn from a Gaussian distribution with mean of u_ϕ and variance of $1/\beta_\phi$:

$$P(\phi_j | u_\phi, \beta_\phi) = N(\phi_j | u_\phi, 1/\beta_\phi) \qquad (2)$$

The model of the ground truth of annotation is established. Since the real annotation of each signal sample is closely related to signal characteristics, a linear regression model is used to simulate the relationship between the ground truth z_i and the feature vector of the ith sample x_i.

$$z_i = w^T x_i + e \qquad (3)$$

In (3), w is the regression coefficient and e is a zero-mean Gaussian noise with variance $1/\alpha$. x_i and w have the same length with $d + 1$, and $x_i = [x_{i1}, x_{i2}, ..., x_{id}, 1]^T$. Note that x_i contains d features because the value of x_{id+1} is always 1. Consequently, the probability density function of z_i is:

$$P(z_i | w, x_i, \alpha) = N(z_i | w^T x_i, 1/\alpha) \qquad (4)$$

Since Gamma distribution is usually used to model a positive continuous value distribution [12], without the prior knowledge, it is assumed that precision parameter β_ϕ of ϕ_j, annotator precision λ^j and precision parameters of z_i are drawn from Gamma distribution:

$$P(\beta_\phi | k_\beta, \vartheta_\beta) = \Gamma(\beta_\phi | k_\beta, \vartheta_\beta) \qquad (5)$$

$$P(\lambda^j | k_\lambda, \vartheta_\lambda) = \Gamma(\lambda^j | k_\lambda, \vartheta_\lambda) \qquad (6)$$

$$P(\alpha | k_\alpha, \vartheta_\alpha) = \Gamma(\alpha | k_\alpha, \vartheta_\alpha) \qquad (7)$$

where Γ denotes a Gamma distribution, k is the shape of distribution and ϑ is the scale of distribution.

Generally, noises such as baseline drift and motion artifacts in poor-quality ECG signals make it hard to detect QRS complexes and to reckon cardiac parameters. Hence, we used four signal quality indices of ECG proposed by Clifford et al. [13] and Li et al. [14] which contains:

- kSQI: the fourth moment (kurtosis) of the distribution, it measures the relative peakedness of the probability distribution with respect to the Gaussian distribution and it can be formulated as $E\{X - \mu\}^4/\sigma^4$.

- sSQI: the third moment (skewness) of the signal which describes symmetrical shape of the distribution and it is defined as $E\{X - \mu\}^3/\sigma^3$.

- pSQI: the relative power in the QRS complex and it can be formulated as $\int_5^{15Hz} P(f)df / \int_5^{40Hz} P(f)df$, where P is the power and f is the frequency.

- basSQI: the relative degree of baseline drift and it is defined as $1 - \int_0^{1Hz} P(f)df / \int_0^{40Hz} P(f)df$.

In addition to the above signal indices, we added two other indices to describe signal quality in this study:

- hSQI: the proportion of high frequency noise in the signal and it is defined as $1 - \int_{40}^{180Hz} P(f)df / \int_5^{180Hz} P(f)df$.

- fSQI: the proportion of non-horizontal segments in the signal, where non-horizontal segment refers to the segment whose amplitude changes between 4 consecutive sampling points are equal to 0 after normalization.

These indices provide the information about the quality and characteristics of ECG signal. Each signal sample to be annotated has a corresponding feature vector, which represents signal situation of current sample. The signal feature with all indices of the ith sample can be presented as:

$$x_i = [kSQI_i, sSQI_i, pSQI_i, basSQI_i, hSQI_i, fSQI_i, 1] \quad (8)$$

In the previous analysis, each annotator is given fixed precision parameters (ϕ_j, λ_j) to measure the accuracy of the annotation, i.e. each annotator behaves consistently throughout all records, but the reality is not. The accuracy is not only related to the performance of the annotator, but also affected by poor quality signal and noise interference. Hence, we added precision adjustment factor of annotator s_i [14] to adjust the algorithm precision throughout the ith records. It is defined as the coincidence between two QRS detection algorithms, i.e. the consistency of QRS detection for the same ECG signal. Suppose that there are K and M QRS position sequences respectively detected by algorithm 1 and 2 for the ith record. If both detected the same QRS waves, then N_{match} plus one. Since the width of QRS wave is between 150ms and 200ms, the error limit of the detection positions is set to 150ms. Finally, according to (9), the precision adjustment factor of annotator for the ith ECG record is calculated.

$$s_i = \frac{N_{match}}{K + M - N_{match}} \qquad (9)$$

After adding the precision adjustment factor s_i to the model, the relationship model between the annotations and the underlying ground truth is shown in (10):

$$P(y_i^j | z_i, \phi_j, \lambda_j) = N(y_i^j | z_i + \phi_j, 1/s_i \lambda_j) \qquad (10)$$

B. Solving Model

The estimated parameters are regarded as random variables that conform to a certain prior probability distribution. The process of observing the sample is the process of converting the prior probability density into the posterior probability density, so that the initial value of the parameter is corrected by sample information. The IBPP model was established. We used Bayesian criterion to obtain the posterior probability of the parameters established in the previous section, then combined Maximum-Posterior (MAP) Estimation and Expectation Maximization (EM) algorithm to iteratively solve the parameters and the underlying ground truth until convergence. The annotations y_i^j and signal characteristic x_i included in (1) are known, observation sample D= $[x_i, y_i^1, y_i^2, ..., y_i^R]_{i=1}^N$, the parameters to be solved are $\theta = \{\phi, \lambda, \omega, \alpha, \beta_\phi\}(\phi = [\phi^1, ..., \phi^R], \lambda = [\lambda^1, ..., \lambda^R])$ and underlying HR value z = $[z_1, ..., z_N]$.

Firstly we obtained the posterior estimation function of the parameter θ according to the Bayesian criterion. The target function of the MAP Estimation is the posterior estimation of the parameter θ related to the observed sample D. The solution formula is:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \quad (11)$$

Since $P(D)$ is a constant that does not depend on θ. The logarithmic posterior estimation of the parameter θ can be regarded as:

$$log\, P(\theta|D) \propto log\, P(D|\theta)P(\theta) \quad (12)$$

In order to solve parameter θ, it can be realized by making the partial derivative of the parameters zero. The estimated ground truth $\hat{z} = [\hat{z}_1, ..., \hat{z}_N]$ was also obtained by this method.

$$\frac{1}{\lambda^j} = \frac{\left[\sum_{i=1}^N s_i\left(y_i^j - \phi^j - \hat{z}_i\right)^2 + \frac{2}{\vartheta_\lambda}\right]}{2(k_\lambda - 1) + N} \quad (13)$$

$$\phi^j = \frac{\left(\sum_{i=1}^N s_i\left(y_i^j - \hat{z}_i\right) + \frac{\mu_\phi \beta_\phi}{\lambda^j}\right)}{\sum_{i=1}^N s_i + \beta_\phi/\lambda^j} \quad (14)$$

$$w = \left(\sum_{i=1}^N x_i x_i^T\right)^{-1} \sum_{i=1}^N \hat{z}_i x_i \quad (15)$$

$$\frac{1}{\beta_\phi} = \frac{\left[\sum_{j=1}^R (\phi^j - \mu_\phi)^2 + \frac{2}{\vartheta_\beta}\right]}{R + 2(k_\beta - 1)} \quad (16)$$

$$\frac{1}{\alpha} = \frac{\left[\sum_{i=1}^N (\hat{z}_i - w^T x_i)^2 + \frac{2}{\vartheta_\alpha}\right]}{N + 2(k_\alpha - 1)} \quad (17)$$

$$\hat{z}_i = \frac{\sum_{j=1}^R s_i \lambda^j \left(y_i^j - \phi^j\right) + \alpha(w^T x_i)}{\sum_{j=1}^R s_i \lambda^j + \alpha} \quad (18)$$

Since there are implicit variable $\hat{z}$ and unknown parameter $\theta = \{\phi, \lambda, \omega, \alpha, \beta_\phi\}$ in the model, (13) to (18) cannot be solved directly. We used the EM algorithm to iteratively solve the MAP estimation of the parameters in the model.

The HR annotation includes two parts: QRS detection and HR calculation. The nine QRS detection algorithms

mentioned in the paper [15] and another improved algorithm were selected, namely: Pan&Tompkins [16], Hamilton_mean [17], Hamilton_median [17], Sixth_power [18], Finite state machine (FSM) [19], U3 transform [20], Difference operation (DOM) [21], Jqrs_improved [22], Optimized Knowledge Based (OKB) [23] and FSM_LFF [19]. According to the step size of 1s, the signal was intercepted every 10 s as a sample, and the HR annotations by jth annotator for the ith signal sample were calculated according to (19).

$$y_i^j = \frac{60}{mRR_i^j} bpm \quad (19)$$

In (19), mRR_i^j represents the median value of the RR intervals annotated by the jth annotator of the ith signal sample. The acquired HR annotation represents the state of the heartbeat change over the past 10 s and is updated once per second.

In addition to MAP and EM algorithm, the concept of selective ensemble of annotation algorithms was also used. Specifically, we selected 3 to N unrepeated algorithms from N ($N \geq 3$) algorithms as an annotator subset each time, and fed the annotations they provide into the IBPP. These algorithms can produce s subsets, which was represented as $\mathcal{B}$, and $\mathcal{B} = \{b_1, b_2, ..., b_s\}$. The EM algorithm required s times iterative operations and output the results of $\hat{z}$、λ、ϕ and RMSE each time. The complete implementation process was shown in Fig. 2 below.

---

for $b_1, b_2, ..., b_s \in \mathcal{B}$,

Input:

Annotation y_i^j provided by annotator subset b_s, i=1,...,N, j=1,...,R (R is the size of the subset b_s)

Feature vector of ECG signal x_i, i=1,...,N

Precision adjustment factor of annotator s_i, i=1,...,N

Prior probability of parameters k_λ、ϑ_λ、k_β、ϑ_β、k_α、ϑ_α

Precision threshold λ_t

Initialise: $\phi, \lambda, w, \alpha, \beta_\phi$

E-step:

Calculate the ground truth $\hat{z}_i$(i=1,...,N) according to the formula (18).

M-step:

Update parameters $\phi, \lambda, w, \alpha, \beta_\phi$ according to the formula (13) to (17).

Repeat E-step and M-step until λ or $\hat{z}$ converges, or the number of iterations reaches an upper limit of 100 times.

Output: $\hat{z}$, λ, ϕ and RMSE

end

Fig. 2. The algorithm flow of the IBPP model.

Generally, root mean square error (RMSE) is used to indicate the square difference between the predicted value and the actual observation. In this paper, the RMSE between the estimated annotations $\hat{z}_i$ and the truth z_i was set as evaluation index. It was calculated as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^N (\hat{z}_i - z_i)^2} \quad (20)$$

III. Experiment and Results

A. Database

We utilized 100 ECG recordings with 10-min or less from the extended training set of the 2014 PhysioNet/CinC challenge [24]. The beats of these ECG recordings were annotated by experts manually, and the calculated HR annotations could be regarded as the referenced ground truth z_i to evaluate algorithms. According to the step size of 1 s, the recordings were intercepted as signal samples every 10 s, and a total of 52886 samples were obtained. It is worth noting that the quality of signal in this dataset is poor, resulting algorithms performing poorly on these signals. Liu *et al.* [15] tested ten widely used QRS detection algorithms on this dataset and found that the accuracies of all algorithms were less than 80%. Therefore, we had the opportunity to utilize the IBPP model to estimate the true HR, test the validity of the model for aggregating continuous labels and appraising the precision of each annotator.

B. Parameter Setting

1) Prior distribution of parameters

First, the precision of the annotator λ, the fitting variance of linear regression for the ECG signal α, and the Gaussian-distribution variance for the annotator bias β_ϕ were mainly determined by the RR interval estimation by Zhu et al. [25],which are shown in Table I. As for the mean μ_ϕ of the Gaussian distribution of the each annotator bias ϕ, we set its initial value as the mean of differences between the HR estimation results for any two algorithms. In the experiment, the model was not sensitive to the values of these parameters when the threshold was determined.

TABLE I. THE VALUES OF THE PRIOR DISTRIBUTION OF THE PARAMETERS IN IBPP MODEL.

Distribution of the parameter	symbol	value
$P(\lambda^j|k_\lambda,\vartheta_\lambda)=\Gamma(\lambda^j|k_\lambda,\vartheta_\lambda)$	k_λ	5.5
	ϑ_λ	1.8
$P(\beta_\phi|k_\beta,\vartheta_\beta)=\Gamma(\beta_\phi|k_\beta,\vartheta_\beta)$	k_β	5
	ϑ_β	0.1
$P(a|k_a,\vartheta_a)=\Gamma(a|k_a,\vartheta_a)$	k_a	3
	ϑ_a	0.6

2) Precision adjustment factor of annotator

In this model, we selected Sixth-power algorithm and OKB algorithm to calculate the precision adjustment factor $s=[s_1,...,s_N]$, and their QRS detection sensitivity and positive rates on the CinC 2014 extended training set were 66.07% and 69.62%, 72.32% and 78.64% respectively [15]. The annotator adjustment factors of 52886 recordings $(s_i, i=1,...,52886)$ in the dataset were obtained, histogram distribution of which are shown in Fig. 3.
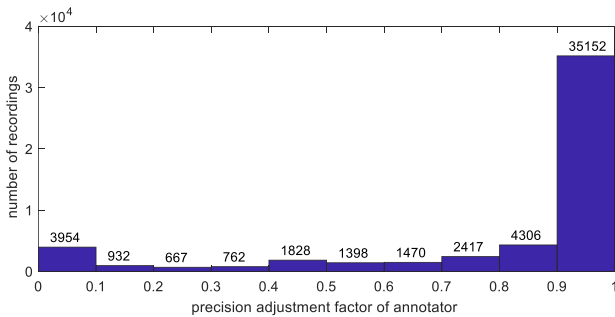
Fig. 3. Distribution histogram of the precision adjustment factor of annotator.

It should be noted that the value of the precision adjustment factor of annotator s cannot be 0, otherwise precision adjustment is meaningless. In Fig. 3, 3954 factors of the recordings were less than 0.1, 2810 factors in which were 0. It indicated that these signals contained too much noises, which made the detection results of two algorithms had a large difference. Therefore, we used the following method to deal with this situation. Assume that the precision adjustment factor for the ith recording is 0, then set it to $t_i/5$, and t_i is defined as normalization of the square sum of the four ECG signals indexes, as shown in equations (21) and (22).

$$t_i = pSQI_i{}^2 + basSQI_i{}^2 + hSQI_i{}^2 + fSQI_i{}^2 \quad (21)$$

$$t_i = \frac{t_i}{max(t_i)} \quad (22)$$

In (21), $pSQI_i, basSQI_i, hSQI_i, fSQI_i$ are characteristic indexes to respectively describe the energy proportion of QRS complex, the degree of baseline drift, the proportion of high-frequency noise, and the proportion of non-linear in the ECG signals. As the same as s_i, the larger the value of t_i, the smaller the noise interference. Because the two methods for calculating the quality of signal and the evaluation criteria were different, it was not guaranteed that t_i was very small when s_i was 0. In order to ensure the relative balance of the substituted values in the original calculation results, $t_i/5$ was required in the range of (0, 0.2).

3) Precision threshold

In order to maximize conditional probability of parameters in each update step, the fusion results will be iterated to the best annotator. However, according to the performance of each annotating algorithm on the CinC 2014 dataset [15], there was obviously no such large precision gap. The simplest and most effective improved method is to set precision threshold. After each iteration of the EM algorithm, the precision greater than the threshold $\lambda^j(j=1,...,R)$ was set to λ^j uniformly and start a new iteration.

In this paper, the Generalized Extreme Value Distribution (GEVD) was used to simulate the distribution of maximum precision (expressed as λ_m). The maximum value of the precision is set as λ_{max}, whose probability density function of the GEVD is:

$$P(\lambda_{max}|k,\mu,\vartheta) = \frac{1}{\vartheta}\left[1 + k\frac{(\lambda_{max}-\mu)}{\vartheta}\right]^{\left(-1-\frac{1}{k}\right)}$$

$$\times exp\left\{-\left[1+k\frac{(\lambda_{max}-\mu)}{\vartheta}\right]^{-\frac{1}{k}}\right\} \quad (23)$$

where k is the shape parameter, ϑ is the proportional parameter, and μ is the positional parameter, these parameters can be obtained by fitting the GEVD. The maximum value of precision λ_m can be obtained by randomly taking m maximums from its prior distribution Gamma (5.5, 1.8). Specifically, we randomly selected m values from the prior distribution of the precision and got the maximum value. After repeating 200 times, a set of precision maximums with size 200 was finally obtained. Using the obtained maximums for GEVD fitting, the probability density function map can be obtained.
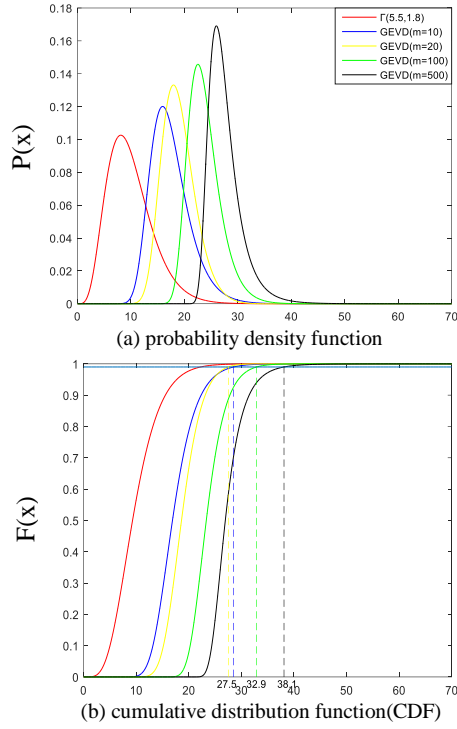
Fig. 4. The GEVD and CDF of Maximum of annotator precision λ: (a) Maximal values were randomly selected from m numbers matching the Gamma (5.5, 1.8) distribution, and the GEVD was fitted. (b) The CDF of generalized extreme value corresponding to different sampling points, the vertical dotted line in the figure indicated the corresponding independent variable when $F(x)$ reached 0.99.

As can be seen in (a) of Fig. 4, the larger the number of m, the larger the maximum value and density, simultaneously the possible value of the precision maximum λ_m was also larger. It can also be clearly seen from (b). The upper bound of annotator precision was obtained by the inverse CDF, that is, the CDF of the precision maximum $F(x)$ was equal to 0.99, and the threshold λ_t was obtained from $P(\lambda < \lambda_t) = 0.99$. Ensure that 99% of the annotation precisions were less than the threshold, and the other 1% were the abnormal extreme values. The annotator subset in this article contained at most 10 annotators, that is, only 10 precision values were from the prior distribution Gamma (5.5, 1.8) at a time. So that m was selected between 10-500 to ensure that most of the maximum values were covered and the interference of extreme outliers was also avoided. Hence the threshold λ_t were determined to be [28, 38]. We set thresholds to 28, 30, 32, 34, 36 respectively. Table II listed the errors with different thresholds. It can be seen that the optimal threshold λ_t was 28. The corresponding RMSE was 11.90 bpm, and the RMSE-Mean and RMSE-Std of were also lower, indicating their overall performance was better and the fluctuation was relatively small.

TABLE II.   THE RMSES OF HR ANNOTATING AT DIFFERENT THRESHOLDS FOR THE IBPP MODEL.

Threshold λ_t	Min-RMSE (bpm)	RMSE- Mean (bpm)	RMSE- Std (bpm)
28	11.90	16.83	3.32
30	11.93	16.88	3.47
32	11.95	16.91	3.60
34	11.98	16.94	3.73
36	12.01	16.98	3.85

C. Results and Analysis

TABLE III.   THE RMSE OF SINGLE ANNOTATING ALGORITHM

Annotator number	Method	RMSE (bpm)	IBPP improved
1	Pan&Tompkins	34.67	58.55%
2	Hamilton_mean	31.04	53.70%
3	Hamilton_median	28.45	49.49%
4	Sixpower	16.66	13.75%
5	FSM	19.79	27.39%
6	U3	34.15	57.92%
7	DOM	32.52	55.81%
8	Jqrs	30.54	52.95%
9	OKB	28.45	49.49%
10	FSM_LFF	35.74	59.79%
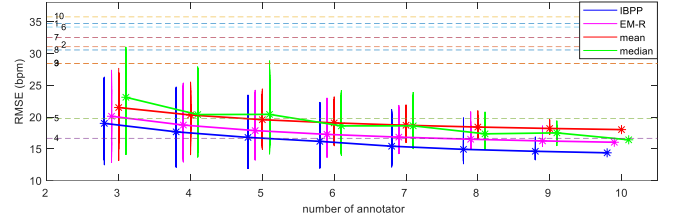
Fig. 5. The HR errors of the IBPP model, EM-R model, mean method and median method with the annotator subsets composed of different number annotators on the dataset. The number 1—10 in the figure respectively corresponds to the single annotating algorithm in Table III.

TABLE IV.   MEAN AND STANDARD DEVIATION OF HR ANNOTATING RMSE FOR FOUR METHODS UNDER DIFFERENT NUMBER OF ANNOTATORS.

Number of labels	IBPP	EM-R	Mean	Median
3	19.00±3.73	20.11±3.96	21.49±3.02	23.10±4.56
4	17.66±3.56	18.73±3.27	20.32±2.44	20.38±3.20
5	16.79±3.27	17.86±2.71	19.58±2.01	20.41±3.54
6	16.17±2.85	17.27±2.24	19.07±1.65	18.60±2.38
7	15.40±2.15	16.83±1.80	18.70±1.33	18.61±2.19
8	14.91±1.56	16.50±1.39	18.42±1.03	17.35±1.41
9	14.60±1.06	16.24±0.97	18.20±0.71	17.50±1.47
10	14.37	16.03	18.02	16.41

Firstly, the paper compared the IBPP model with other continuous-valued fusion methods by CinC 2014 extended training set. 3-10 annotators were randomly selected from 10 annotators and a total of 968 annotator subsets were obtained. All the annotator subsets were input into four methods. The results were grouped according to the identical number of annotators in the subset, and mean and standard deviation of RMSE within the group were calculated and listed in Table IV. As shown in Fig. 5, the IBPP model outperformed the EM-R model, mean and median method under the same number of annotators. As the number of fusing annotators increased, the RMSE mean decreased. The highest improvement of the IBPP over the EM-R, mean and median annotation fusing methods were 10.36%, 20.26% and 12.43% respectively when considering all 10 annotations and the IBPP model got the minimum RMSE of 14.37 bpm.

Then the results of the IBPP model with the 10 single annotating algorithms were also compared. Table III listed the

RMSEs of 10 single algorithms on the CinC 2014 extended training set. The highest-precision annotator was the Sixth-power algorithm (16.66 bpm), and the worst was the FSM_LFF algorithm (35.74 bpm). The lowest RMSE (14.37 bpm) from IBPP using 10 annotators had an improved error rate of 13.75% and 59.79% when compared with the optimal and worst single annotator. Fig. 5 also compared the RMSEs of the four methods with 10 single algorithms. It can be seen that when the number of the fusion annotators was greater than 5, the IBPP model outperformed than the best annotator Sixth-power. As for all the annotator subsets, the IBPP model outperformed than other nine single algorithms.

In addition, when HR annotations for 52886 ECG samples annotated by 10 annotators were input into the IBPP model, it took about 0.1 seconds for EM algorithm to run on the 3.2 GHz Intel Core processor using MATLAB software until convergence.

IV. DISCUSSION AND CONCLUSION

This paper proposed an unsupervised HR estimation model combined with Bayesian prior probability here annotators' accurate and the ground truth are not available. The model added prior distributions of annotator's precision parameters, ECG linear regression fitting noise parameters, and annotator bias parameters, then introduced MAP and EM algorithm to solve the model. The results showed that the IBPP model had different degrees of improvement in HR estimation accuracy. Even if the different annotators subsets, the error of IBPP model was still the smallest in the four methods. Stability and superiority of the IBPP model were demonstrated. Importantly, the IBPP model does guarantee a better fusion result than each single annotator without any prior knowledge of who or what is the best annotator. The proposed model can be used in the fusion of manual or algorithmic annotation results.

The novelty of the proposed method should be mentioned. On one hand, the IBPP model provides an estimation of continuous-valued annotations in an unsupervised Bayesian framework to infer the underlying ground truth and annotators' precisions, no referenced data is required to train the model parameters. On the other hand, it includes an annotator precision adjustment factor. Previous studies assumed the annotator performance does not depend on signal, which is not true in practice. Because the performance of the same annotator may be affected by signal noises, the annotator precision adjustment factor satisfies this actual situation. We adjust the precision of the annotator by considering whether the quality of ECG signal can ensure the identification of the characteristic waveform. However, the proposed model needs to be improved for application in the future work.

REFERENCES

[1] World Health Organization. Cardiovascular Diseases (CVDs) [EB/OL]. Available: http://www.who.int/mediacentre/factsheets/fs317/en/

[2] C. Y. Liu, X. Y. Zhang, L. N. Zhao, F. F. Liu, X. W. Chen, et al., "Signal quality assessment and lightweight qrs detection for wearable ecg smartvest system," IEEE. Internet. Things, vol. 6, pp. 1363-1374, 2019.

[3] R. M. Laukkanen and P. K. Virtanen, "Heart rate monitors: State of the art," J. Sport. Sci, vol. 16, pp. 3-7, 1998.

[4] C. Y. Liu, M. C. Yang, J. N. Di, Y. T. Xing, Y. W. Li, et al., "Wearable ECG: History, key technologies and future challenges," Chinese Journal of Biomedical Engineering, vol. 38, pp. 641-652, 2019.

[5] PhysioNet. MIT-BIH Arrhythmia Database [EB/OL]. Available: https://www.physionet.org/physiobank/database/mitdb/

[6] T. T. Zhu, A. E. W. Johnson, J. Behar, and G. D. Clifford, "Crowd-sourced annotation of ECG signals using contextual information," Ann. Biomed. Eng, vol. 42, pp. 871-884, 2013.

[7] A. P. Dawid and A. M. Skene, "Maximum likelihood estimation of observer error-rates using the EM algorithm," J. Roy. Stat. Soc, vol. 28, pp. 20-28, 1979.

[8] B. Carpenter, "Multilevel bayesian models of categorical data annotation," unpublished.

[9] V. C. Raykar, S. P. Yu, L. H. Zhao, G. H. Valadez, C. Florin, et al., "Learning from crowds," J. Mach. Learn. Res, vol. 11, pp. 1297-1322, 2010.

[10] P. Welinder and P. Perona, "Online crowdsourcing: Rating annotators and obtaining cost-effective labels," IEEE CVPR 2010, pp. 25-32, 2010.

[11] Y. T. Xie, J. Q. Li, T. T. Zhu, and C. Y. Liu, "Continuous-valued annotations aggregation for heart rate detection," IEEE. Access, vol. 7, pp. 37664-37671, 2019.

[12] C.M. Bishop, Pattern Recognition and Machine Learning, 1st ed., New York: Springer, 2006, pp. 271-350.

[13] G. D. Clifford, J. Behar, Q. Li, and I. Rezek, "Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms," Physiol. Meas, vol. 33, pp. 1419-1433, 2012.

[14] Q. Li, R. G. Mark, and G. D. Clifford, "Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter," Physiol. Meas, vol. 29, pp. 15-32, 2008.

[15] F. F. Liu, C. Y. Liu, X. G. Jiang, Z. M. Zhang, Y. T. Zhang, et al., "Performance analysis of ten common QRS detectors on different ECG application cases," J. Healthc. Eng, vol. 2018, pp.1-8, 2018.

[16] J. Pan, and W. J. Tompkins, "A real-time QRS detection algorithm," IEEE. T. Bio-med. Eng, vol. 32, pp. 230-236, 1985.

[17] P. S. Hamilton and W. J. Tompkins, "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database," IEEE. T. Bio-med. Eng, vol. 33, pp. 1157-1165, 1986.

[18] A. K. Dohare, V. Kumar, and R. Kumar, "An efficient new method for the detection of QRS in electrocardiogram," Comput. Electr. Eng, vol. 40, pp. 1717-1730, 2014.

[19] R. Gutierrez-Rivas, J. J. Garcia, W. P. Marnane, and A. Hernandez, "Novel real-time low-complexity QRS complex detector based on adaptive thresholding," IEEE. Sens. J, vol. 15, pp. 6036-6043, 2015.

[20] M. Paoletti and C. Marchesi, "Discovering dangerous patterns in long-term ambulatory ECG recordings using a fast QRS detection algorithm and explorative data analysis," Comput. Meth. Prog. Bio, vol. 82, pp. 20-30, 2006.

[21] Y. C. Yeh and W. J. Wang, "QRS complexes detection for ECG signal: the Difference operation method," Comput. Meth. Prog. Bio, vol. 91, pp. 245-254, 2008.

[22] J. Behar, J. Oster, and G. D. Clifford, "Non-invasive FECG extraction from a set of abdominal sensors," CINC 2013, vol. 40, pp. 297-300, 2013.

[23] M. Elgendi, "Fast QRS detection with an optimized knowledge-based method: Evaluation on 11 standard ECG databases," Plos One, vol. 8, pp. 73557, 2013.

[24] G. Moody, B. Moody, and I. Silva, "Robust detection of heart beats in multimodal data: The physioNet/computing in cardiology challenge 2014," CINC 2014, vol. 41, pp. 549-552, 2015.

[25] T. T. Zhu, N. Dunkley, J. Behar, D. A. Clifton, and G. D. Clifford, "Fusing continuous-valued medical labels using a Bayesian model," Ann. Biomed. Eng, vol. 43, pp. 2892-2902, 2015

Continuous Estimation of Left Ventricular Hemodynamic Parameters Based on Heart Sound and PPG Signals Using Deep Neural Network

Tengfei Feng
School of Biomedical Engineering
Dalian University of Technology
Dalian, China
tengfei.feng@mail.dlut.edu.cn

Hong Tang*
School of Biomedical Engineering
Dalian University of Technology
Dalian, China
tanghong@dlut.edu.cn

Miao Wang
School of Biomedical Engineering
Dalian University of Technology
Dalian, China
wangmiao9248@mail.dlut.edu.cn

Chi Zhang
School of Biomedical Engineering
Dalian University of Technology
Dalian, China
chizhang@dlut.edu.cn

Hongkai Wang
School of Biomedical Engineering
Dalian University of Technology
Dalian, China
wang.hongkai@dlut.edu.cn

Fengyu Cong
School of Biomedical Engineering
Dalian University of Technology
Dalian, China
cong@dlut.edu.cn

Abstract—**Continuous estimation of left ventricular hemodynamic parameters is helpful to early diagnosis of cardiovascular diseases. Current non-invasive methods are somewhat inconvenient to monitor these parameters. Here, a deep neural network is built to noninvasively estimate left ventricular systolic pressure (LVSP), left ventricular diastolic pressure (LVDP), maximum rate of left ventricular pressure rise (+ dp/dt(max)) and minimum rate of left ventricular pressure drop (- dp/dt(min)) based on heart sound and PPG signals. The model consists of residual network and bi-directional recurrent neural network. Performance is evaluated on 2 beagle dogs' experiment data with large ranges induced by epinephrine. Mean absolute errors and standard deviations between the estimated and the measured LVSP, LVDP, + dp/dt(max) and - dp/dt(min) are 7.23±8.33 mmHg, 2.12±3.0 mmHg, 298±406 mmHg/s, and 172±386 mmHg/s, respectively. The average correlation coefficients for LVSP, LVDP, + dp/dt(max) and - dp/dt(min) are 0.94, 0.86, 0.95 and 0.92. The results show that accurate intraventricular hemodynamic parameters can be achieved by non-invasive heart sound and PPG signals with deep neural networks. This technique suggests an easy way for real-time monitoring of intraventricular hemodynamics.**

Keywords—left ventricular systolic pressure, left ventricular diastolic pressure, maximum rate of left ventricular pressure rise, minimum rate of left ventricular drop, deep learning, heart sound signal, PPG signal

I. INTRODUCTION

Monitoring hemodynamic parameters in left ventricle is essential to the detection and treatment of heart failure, myocardial ischemia, and other cardiovascular diseases (CVDs). Previous studies illustrated that hemodynamic changes could be observed before occurrence of symptoms and continuous monitoring of hemodynamic parameter is benefit to detect early hints of dysfunction [1, 2].

Heart sound (HS) and photoplethysmograph (PPG) signals, which are forceful assistants for inspecting CVDs, can be implemented in low-cost. Previous studies have shown that there are strong relations between HS, PPG signals and blood pressure (BP). In 1960s, Shah et al [3] and Sakamoto et al [4] had discovered that there were linear relationships between the rising rate of left ventricular BP and the amplitude of the

first heart sound. Reference [5–7] extracted features of HS signals from different domains to estimate SP and DP, which achieved high correlation coefficients. Previous studies [8–11] estimated BP parameters based on PPG signals, showing good performance and meeting the Association for the Advancement of Medical Instrumentation (AAMI) standard. An integrated chest wearable apparatus for continuous blood pressure estimation using PPG and phonocardiogram (PCG) was introduced in the work of Marzorati et al [12], which showed a promising application of systolic pressure (SP), diastolic pressure (DP) monitoring algorithms based on PPG and PCG signals.

After investigating previous literatures, we found (1) intraventricular BP values are seldom estimated, (2) only one hemodynamic parameter was estimated by extracted multi-domain features in most studies, (3) short ranges of BP values were considered in previous studies.

In this work, an end-to-end deep neural network, whose inputs are the segments of HS and PPG signals of one cardiac cycle, is established to determinate four left ventricular hemodynamic parameters, left ventricular systolic pressure (LVSP), left ventricular diastolic pressure (LVDP), maximum rate of left ventricular pressure rise (+ dp/dt(max)) and minimum rate of left ventricular pressure drop (– dp/dt(min)). The data was measured from two beagle dogs who had a large range of hemodynamics by injecting different doses of epinephrine. The HS signal, BP, PPG signal, and electrocardiograph (ECG) are simultaneously recorded in the whole procedure of the experiment. The collected data contains over 11,000 cardiac cycles, which are not bad for training the deep neural network. The performance reveals that the left intraventricular hemodynamic parameters can be
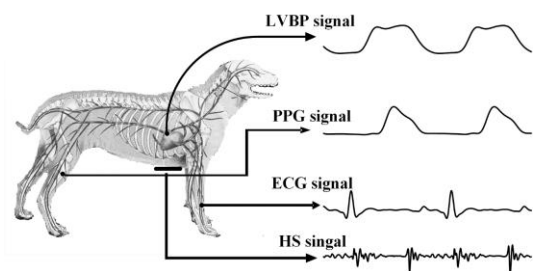
Fig. 1.   Data collection scheme of the experiment

*Hong Tang is the corresponding author. (email: tanghong@dlut.edu.cn).
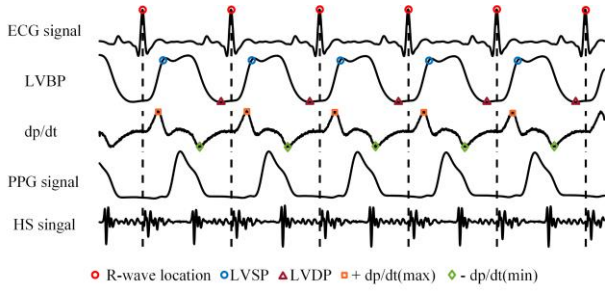
Fig. 2. Scheme of splitting the cardiac cycles according to the location of R-waves

estimated with HS and PPG signals. There would be promising applications in non-invasive intraventricular hemodynamic parameters monitoring.

II. MATERIALS AND METHODS

A. Data Acquisition and Preprocessing

The experiment was approved by the Animal Care Committee of Chongqing Medical University and two healthy beagle dogs weighing 9-10kg are involved. The dogs laid down calmly in the supine position by anesthetizing xylazine (0.2 ml/kg) at first. A catheter filled with heparinized solution (500 units/ml) was inserted into the left ventricle via the carotid artery. Then, different doses of epinephrine (0.5 g/kg, 1 g/kg, or 2 g/kg) were injected to the dogs via a route formed by an intravenous infusion of 0.9% saline. Data was collected from 10 seconds before injection epinephrine to BP values back to normal. The procedure was repeated 3-5 times. As the Fig. 1 shows, ECG, HS, PPG, and BP signals were recorded simultaneously in the whole procedure. The BP signal was recorded by a pressure transducer (MLT0699, ADInstruments, Australia) which was connected to the inserted catheter and calibrated at standard atmospheric pressure. The ECG electrodes were placed on both of the dog's forelimbs as lead I. A microphone transducer (MLT201, ADInstruments, Australia) was placed at the apex of the heart to record external HSs. PPG was recorded noninvasively by a photoplethysmogram sensor (MLT1020FC, ADInstruments, Australia) which was affixed to the femoral artery. The digital sampling frequency of all signals is 1KHz. In the end, 29 records were obtained. The details of the records are shown in Table I.

R-waves of ECG were identified by Pan &Tompkins (P&T) algorithm [13]. As shown in Fig. 2, all the signals are cut into cardiac cycles in accordance with R-wave locations. The BP values at the end of systolic stroke are detected as LVSP. LVDP was achieved by detecting the minimum values of BP in diastolic phases. The maximum and minimum values of BP's derivatives were extracted as + dp/dt(max) and – dp/dt(min) respectively. Eventually, 7529 and 3935 cardiac cycles were collected after preprocessing for subject 1 and subject 2 respectively.

Fig. 3. Structure of the end-to-end model for estimating hemodynamic parameters. The inputs are HS and PPG signals of one cardiac cycle, which are zero padded to 1000 samples, and outputs are the four hemodynamic parameters, which are scaled by weights in order to yield equal contribution to the loss function.

B. Network Architecture for Hemodynamic Parameters Estimation

The estimation of hemodynamic parameters could be seen as time-series regression problem. Recent years, deep neural network has displayed outstanding performance for regression tasks in many fields. As illustrated in Fig. 3, an end-to-end model formed by residual network and bidirectional recurrent neural network (Bi-RNN) is built to learn features from HS and PPG signals and estimate the hemodynamic parameters.

The inputs for the model are HS and PPG signals of one cardiac cycle, which have been padded to 1000 samples as the time duration of one cycle are different due to heart rate variability (HRV). After the input layer, residual network with six residual blocks is used to learn local features. The residual blocks' architecture is reference to the model used in [14]. The component for residual blocks are clearly shown in Fig. 3. The probability for dropout layer is set to be 0.2. Some parameters of residual blocks are displayed in Table II. the pooling size and the stride of the second convolutional layer in the 1st, 3rd, 5th residual blocks are 2, which means that the length of feature vector reduces twice after getting through these blocks. Therefore, the output length of this part is 125(1/8 of the input length).

Next to the residual network, bidirectional recurrent neural network (Bi-RNN) module is used to learn the context relations between the different local features as HS and PPG signals (also the learned feature vectors) are time series. Gate Recurrent Unit (GRU) was selected as implementation of Bi-RNN. Here, the unit number is set to 32, so the feature vector length of this part is 64.

Finally, the extracted features are input to a dense layer and output the four estimated hemodynamic parameters values. Then, the estimated values are compared with the reference hemodynamic parameters using the loss function. As LVSP, LVDP, + dp/dt(max) and – dp/dt(min) are clustered at 50-300, -30-20, 0-10000 and 1000-6000 respectively,

	Num. of record	Num. of cardiac cycles	LVSP (mmHg) (Min - Max)	LVDP (mmHg) (Min - Max)	+ dp/dt(max) (mmHg/s) (Min - Max)	– dp/dt(min) (mmHg/s) (Min - Max)
Subject 1	15	7529	77 - 268	-47 - 10	107 - 9181	-5896 - -70
Subject 2	14	3935	122 - 272	-35 - 18	106 - 8198	-4973 - -175
Total	29	11,464	77 - 272	-47 - 18	106 - 9181	-5896 - -70

TABLE I. DETAILS OF THE MEASURED DATA.

TABLE II.   DETAILS OF THE PARAMETERS IN RESIDUAL BLOCKS

Residual blocks	Num. of Conv	Kernel length	Stride of Conv	Pooling size
1st	#1	16	1	2
	#2	16	2	
2nd	#1	16	1	1
	#2	16	1	
3rd	#1	16	1	2
	#2	16	2	
4th	#1	16	1	1
	#2	16	1	
5th	#1	16	1	2
	#2	16	2	
6th	#1	16	1	1
	#2	16	1	

LVSP/50, LVDP/10, +dp/dt(max)/2000 and − dp/dt(min)/1000 are used as the reference values so that each parameter has same contributes to the loss function. Then, the parameters of the model are updated via back propagation. And the initial learning rate for gradient descent is set to be 0.001.

C. 5-fold Cross-validation Scheme and Evaluate Indicators

The records of one subject are randomly divided into 5 subsets. Then, 5-fold cross-validation are evaluated. For instance, 15 records of subject 1 are divided into 5 equal subsets, each of which has 3 records. Then, using the first 3 records as test data, the other 12 records are split into training data and validation set by 3:1. The model is training 200 epochs on training data and evaluate on validation data. The model with least loss on validation data is chosen to evaluate performance on test data. Repeating the procedure on the remained 4 subsets, all the data are selected as test data once.

In this study, Pearson correlation coefficient (CC), mean error(ME), mean absolute error (MAE) and standard deviation (SD) are used to evaluate performance. Their formulas are follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{1}$$
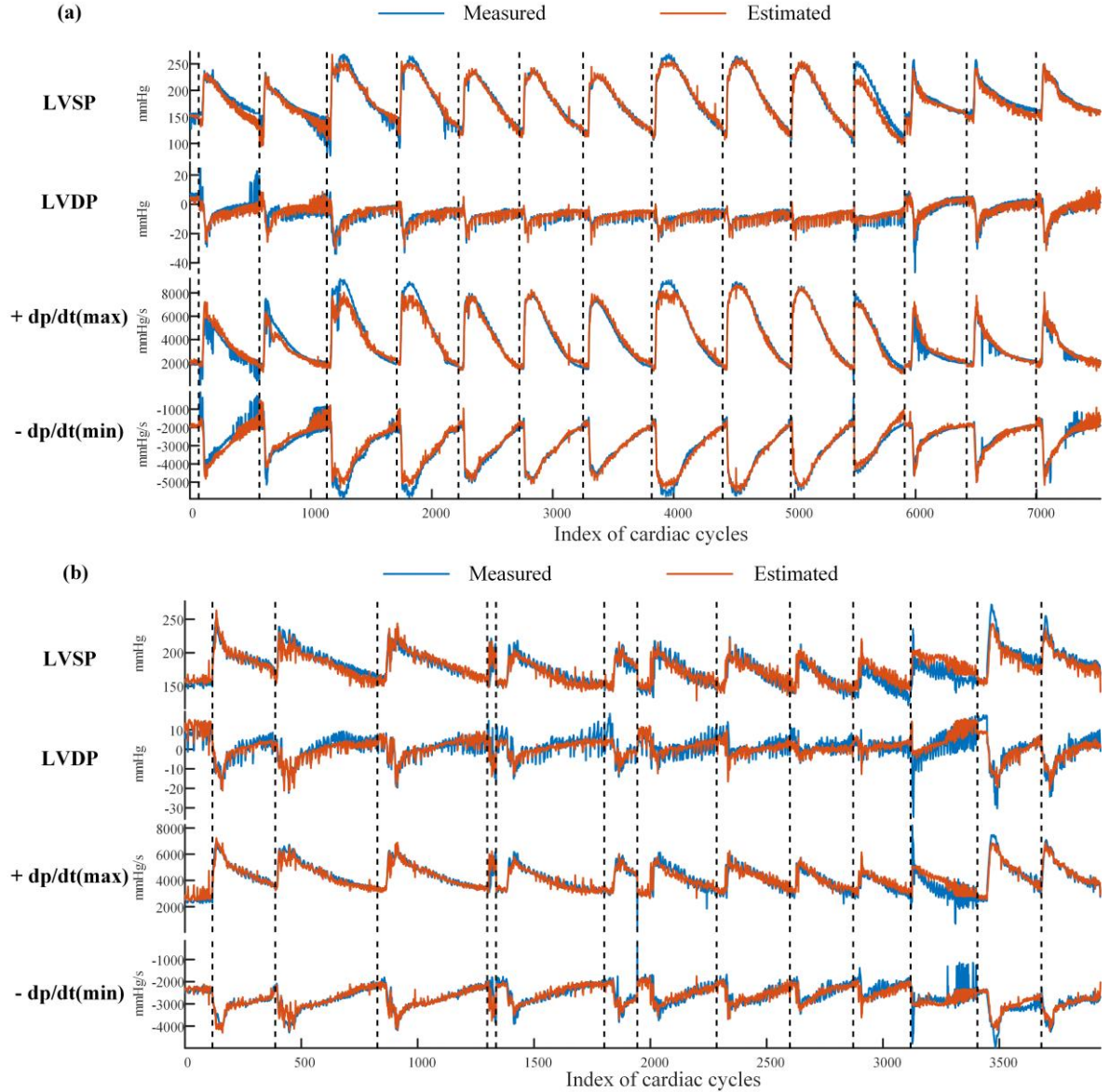
Fig. 4.   Estimated hemodynamic parameters and the measured. (a) 15 records of subject 1. (b) 14 records of subject 2

$$ME = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i) \qquad (2)$$

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\hat{y}_i - y_i - ME)^2} \qquad (3)$$

$$CC = \frac{\sum_{i=1}^{N}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2}} \qquad (4)$$

where $\hat{y}$ is the estimated value, and y is the measured value. $\bar{\hat{y}}$ and $\bar{y}$ are the average values. N is number of cardiac cycles involved in.

III. RESULTS

As shown in Fig. 4, the estimated hemodynamic parameters by the deep neural network are compared with the measured. It can be observed clearly that the estimated follows the measured values very well for all parameters. As correlation analysis displayed in Fig. 5, there are strong correlation between estimated and measured values. The average of 5-fold cross-validation evaluate indicators are illustrated detailedly in Table III. All evaluate indicators' results are presented in the following order: LVSP, LVDP, + dp/dt(max), and – dp/dt(max). For subject 1, The MEs $\pm$ SDs are -3.41 $\pm$ 6.69 mmHg, -0.1 $\pm$ 2.42 mmHg, -39 $\pm$ 428 mmHg/s and 59 $\pm$ 226 mmHg/s. The CCs are 0.984, 0.916, 0.979 and 0.972. For subject 2, the MEs $\pm$ SDs are -0.12 $\pm$ 9.97 mmHg, -0.4 $\pm$ 3.59 mmHg, 37 $\pm$ 383 mmHg/s and -0.9 $\pm$ 546 mmHg/s, the CCs are 0.897, 0.812, 0.922 and 0.873. The average MAEs are 7.23 mmHg, 2.12mmHg, 298 mmHg/s, and 172 mmHg/s, and the relative errors are less than 5% of the corresponding range. This results demonstrates that there are strong links between the HS and PPG signals and the intraventricular hemodynamics. And the relationships highly generalized in records from one subject.

IV. DISCUSSIONS

Lots of previous works have been studied on estimation of SP based on heart sound and/or PPG signals. It is wise because SP is one of the most important hemodynamic parameters and
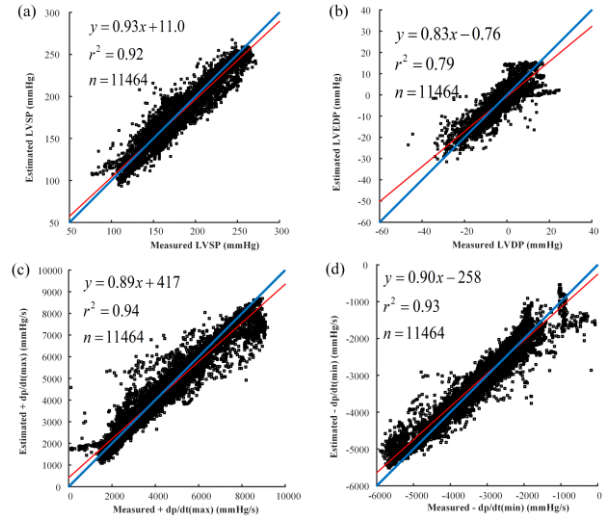
Fig 5.   Correlation analysis between the estimated and the measured. (a) LVSP, (b) LVDP, (c) + dp/dt(max), (d) – dp/dt(min)

the afterload of left ventricular wall at the end of blood ejection can be reflected by SP. Many abnormal states could be mirrored by LVSP values, like compliance of arteries, peripheral vascular resistance and others. LVSP can be measured indirectly by cuff BP at artery is another main reason for the popular study of LVSP.

However, LVDP, + dp/dt(max) and – dp/dt(min), which are difficult to be measured in clinical conditions are seldom concerned and estimated noninvasively. LVDP is an index of intraventricular preload, like the initial load for left ventricle before contraction. And there are relations between LVDP value and venous return volume. For instance, return volume will increase due to aortic stenosis, and then LVDP increased. Inversely, mitral stenosis could lead to decreasing of return volume, then LVDP reduces. + dp/dt(max) and – dp/dt(min) are important indexes for evaluating contractility function of ventricle. In this study, the deep neural network shows an advantage of building the links between the four hemodynamic parameters and HS and PPG signals. The model suggests a way to beat-by-beat estimate four parameters simultaneously and can be used to much wider hemodynamic range than former studies. It hopes to be implemented in health monitoring.

TABLE III.   EVALUATION PERFORMANCE OF 5-FOLD CROSS-VALIDATION

	Indicators	LVSP (mmHg)	LVDP (mmHg)	+ dp/dt(max) (mmHg/s)	- dp/dt(min) (mmHg/s)
Subject 1	ME	-3.41	-0.1	-39	59
	MAE	6.22	1.54	329	175
	SD	6.69	2.42	428	226
	CC	0.984	0.916	0.979	0.972
	P	<<0.001	<<0.001	<<0.001	<<0.001
	95% CI for CC	0.982-0.985	0.908-0.924	0.977-0.981	0.969-0.975
Subject 2	ME	-0.12	-0.4	37	-0.9
	MAE	8.23	2.77	267	169
	SD	9.97	3.59	383	546
	CC	0.897	0.812	0.922	0.873
	P	<<0.001	<<0.001	<<0.001	<<0.001
	95% CI for CC	0.883-0.910	0.787-0.834	0.911-0.932	0.855-0.888

	Signal	BP	Method	SP (mmHg)	Performance	Relative scale	Output
Tang et al [6]	HS	LVSP	Multi domain features + SVM[a]	77-272	CC:0.92 MAE:6.86mmHg SD:8.96mmHg	MAE:3.5% SD:4.6%	Beat-by-beat
Peng et al [5]	HS	Finger cuff BP	Spectrum of S2 + SVM	90-140	CC:0.707 MAE:4.34mmHg SD:6.12mmHg	MAE:8.6% SD:12.2%	Beat-by-beat
Kapur et al [7]	HS	Intra-arterial BP	Characteristics of S1 and S2 + ANN[b]	58-173	CC:0.964 RMSE[c]:7.305mmHg SD:7mmHg	RMSE:6.3% SD:6.1%	Every 10s
Esmaelpoor et al [15]	PPG	Invasive BP	Deep neural network	80-180	MAE:3.97mmHg SD:5.55mmHg	MAE:4.0% SD:5.6%	Beat-by-beat
Yan et al [9]	PPG+ECG	Arterial BP	Deep CNN[d]	80-180	CC: 0.966 MAE:3.09mmHg SD:2.76mmHg	MAE:3.1% SD:2.8%	Every 10s
This work	HS +PPG	LVSP	Deep learning	77-272	CC: 0.982 MAE: 4.34mmHg SD: 6.37mmHg	MAE:2.2% SD:3.3%	Beat-by-beat

[a.] SVM, support vector machine.

[b.] ANN, artificial neural network.

[c.] RMSE, root mean square error.

[d.] CNN, convolutional neural network.

The results showed that the estimation performance of LVDP seems worse than the other parameters. To the authors' knowledge, this degradation is not due to the model but the weak physiological relation. Previous physiological studies revealed that the physiological link between LVDP and PCG, PPG was somewhat weaker than those links of the other three parameters. However, we think degraded performance of LVDP is applicable. We can see that the estimated LVDP has error of 2.12±3.0 mmHg. This accuracy could provide useful hemodynamic information for diagnosis.

For evaluating performance of the estimation method, different cross validation schemes can be designed in different aspects. As mentioned in Section II, 29 records with 11,464 cardiac cycles have been collected. A validation scheme for separating the data into training and testing data is the scheme introduced in Section IIC and we call this scheme as scheme I. Another common way in previous studies is to divide the data according to cardiac cycles whatever which record and which subject the cycles are from. That is, the cardiac cycles are equally and randomly split into five subsets to carry 5-fold cross-validation. This way is called as Scheme 0 in this study. The structure of the network is same as shown in Fig. 3. With Scheme 0, the average of 5-fold cross-validation MEs±SDs are 0.08±6.37 mmHg, -0.16±2.09 mmHg, 82.8±275.2 mmHg/s, and -3.2±170 mmHg/s for LVSP, LVDP, + dp/dt(max), and – dp/dt(min), the average MAEs are 4.34 mmHg, 1.34 mmHg, 203 mmHg/s, and 106 mmHg/s.

Akin works on estimation of BP parameters evaluating like Scheme 0 are investigated and compared with the result we have achieved, as shown in Table IV. Reference [6] extracted multi features from HS signal to estimate LVSP with SVM and analyzed the relationships between SP and the features. Each record was divided into ten parts to evaluate the performance via 10-fold cross validation. 6.86 mmHg, 8.96mmHg, and 0.92 were achieved for average MAE, SD, and CC, respectively. Reference [5] used the Fourier spectrum of the second heart sound to estimate finger cuff BP with regression model. To achieved variational BP, a cold-pressor experiment was carried on 32 healthy subjects. Cardiac cycles of each subject were randomly divided to do 10-fold cross-validation. The average values of SP were 0.707, 4.339mmHg,

and 6.121mmHg for CC, MAE, and SD. An artificial neural network was used to determinate inter-arterial (lower limbs) SP and DP using features of HS in the work of Kapur et al [7]. Data was collected from 25 children with artificial heart valves. One output was given for every 10-second signal. The performance of SP was 7.31 mmHg for RMSE and 7 mmHg for SD. A multistage deep neural network was proposed in [15] to estimate arterial BP and DP based on PPG signals. The records of 200 patients contains 51,884 heart beats and each record was divided (70% for train, 10% for validation, and 20% for test). The range for SP is 80-180 and the results of SP are 3.97 mmHg for MAE and 5.55 mmHg for SD. A deep convolutional neural network was developed in [9] to estimate arterial SP and DP. 604 subjects were selected from the dataset and each subject contains 190-200 segments (10-second). When splitting randomly all subjects' segments, the results of SP were 0.966, 3.09 mmHg, and 2.76 mmHg for CC, MAE, and SD respectively. Then, 604 subjects were randomly split into ten parts to carry 10-fold cross validation. For SP, the results were 0.224, 12.49 mmHg, and 9.43 mmHg for CC, MAE, and SD.

We could find that the model with validation Scheme 0 has a better MAE and SD (4.34 and 6.37). Think the two schemes over, we could get conclusions. (1) The results with Scheme 0 are better, however, it is unrealistic for application in clinical because the data was included in the training phase to some extent. Training a model in this way is impossible in practical engineering. (2) With Scheme I, the good performance shows the high relations among HS, PPG and the four hemodynamic parameters and it is possible to training a model in this way to monitor hemodynamic parameters in this way for a certain subject.

We investigated the over-fitting problem of the model by analyzing performance difference between the training and the testing. The computer simulation showed that the MAE performances of the four parameters in training phase were 5mmHg, 1.38 mmHg, 184 mmHg/s and 116 mmHg/s. However, the performances in testing phase were 7.23 mmHg, 2.12mmHg, 298 mmHg/s, and 172 mmHg/s. If the training performance was much higher than testing performance, it can be inferred that the over-fitting problem occurred. However,

the performances had little difference from the training to the testing. Therefore, the authors concluded that the proposed model had little over-fitting problem. Moreover, the computational complexity of the proposed model was evaluated by computation simulations. Once the model was trained, 6.4 ms elapsed in average to output the estimation with input data of one cardiac cycle. The time consumption is short enough to make beat-by-beat response. Therefore, the computational complexity is not a big problem in this application.

In short, the proposed model has advantages. (1) The model outputted the hemodynamic parameters in manner of beat-by-beat. In another words, the parameters could be obtained in almost real time. (2) The model outputted four hemodynamic parameters simultaneously, i.e., LVSP, LVDP, +dp/dt(max) and –dp/dt(min). The parameters are critically important in assessment of heart function. (3) The parameters estimated by the model are accurate. The computer results showed that the relative errors were less than 5%. The accuracy is potentially applicable in clinical diagnosis.

V. CONCLUSIONS

In this study, we developed a deep neural network to estimate simultaneously four hemodynamic parameters based on noninvasive HS and PPG signals. The model consisted of residual network and Bi-RNN module, which extracted the local features and context relationships. The evaluation performance demonstrated that the neural network could construct the beat-by-beat relations between the four left ventricular hemodynamic parameters and the collected HS, PPG signals. This study states that the left intraventricular hemodynamic parameters, which are difficult to measure clinically, could be beat-by-beat estimated by conveniently obtained HS and PPG signals. There would be potential applications of the proposed method in early diagnosis of hemodynamic related diseases.

REFERENCES

[1] M. R. Zile et al., "Transition from chronic compensated to acute decompensated heart failure: pathophysiological insights obtained from continuous monitoring of intracardiac pressures," Circulation, vol. 118, no. 14, pp. 1433–1441, Sep. 2008.

[2] T. Mondritzki et al., "Remote left ventricular hemodynamic monitoring using a novel intracardiac sensor," Circ. Cardiovasc. Interv., vol. 11, no. 5, May 2018.

[3] P. M. Shah, M. Mori, D. M. Maccanon, and A. A. Luisada, "Hemodynamic correlates of the various components of the first heart sound," Circ. Res., vol. 12, no. 4, pp. 386–392, May 1963.

[4] T. Sakamoto, R. Kusukawa, D. M. Maccanon, and A. A. Luisada, "Hemodynamic determinations of the amplitude of the first heart sound," Am. J. Cardiol., vol. 15, no. 1, pp. 143–143, Feb. 1966.

[5] R.-C. Peng, W.-R. Yan, N.-L. Zhang, W.-H. Lin, X.-L. Zhou, and Y.-T. Zhang, "Cuffless and continuous blood pressure estimation from the heart sound signals," Sensors, vol. 15, no. 9, pp. 23653–23666, Sep. 2015.

[6] H. Tang, J. Zhang, H. Chen, A. Mondal, and Y. Park, "A non-invasive approach to investigation of ventricular blood pressure using cardiac sound features," Physiol. Meas., vol. 38, no. 2, pp. 289–309, Feb. 2017.

[7] G. Kapur et al., "Noninvasive determination of blood pressure by heart sound analysis compared with intra-arterial monitoring in critically ill children—a pilot study of a novel approach," Pediatr. Crit. Care Med., vol. 20, no. 9, pp. 809–816, Sep. 2019.

[8] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," IEEE Trans. Biomed. Eng., vol. 64, no. 4, pp. 859–869, Apr. 2017.

[9] C. Yan et al., "Novel deep convolutional neural network for cuff-less blood pressure measurement using ECG and PPG signals," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, Jul. 2019, pp. 1917–1920.

[10] S. G. Khalid, J. Zhang, F. Chen, and D. Zheng, "Blood pressure estimation using photoplethysmography only: comparison between different machine learning approaches," J. Healthc. Eng., vol. 2018, pp. 1–13, Oct. 2018.

[11] X. Ding, B. P. Yan, Y.-T. Zhang, J. Liu, N. Zhao, and H. K. Tsang, "Pulse transit time based continuous cuffless blood pressure estimation: a new extension and a comprehensive evaluation," Sci. Rep., vol. 7, no. 1, pp. 11554, Dec. 2017.

[12] D. Marzorati, D. Bovio, C. Salito, L. Mainardi, and P. Cerveri, "Chest wearable apparatus for cuffless continuous blood pressure measurements based on PPG and PCG signals," IEEE Access, vol. 8, pp. 55424–55437, Mar. 2020.

[13] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," IEEE Trans. Biomed. Eng., vol. BME-32, no. 3, pp. 230–236, Mar. 1985.

[14] A. Y. Hannun et al., "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," Nat. Med., vol. 25, no. 1, pp. 65–69, Jan. 2019.

[15] J. Esmaelpoor, M. H. Moradi, and A. Kadkhodamohammadi, "A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals," Comput. Biol. Med., vol. 120, pp. 103719, May 2020.

Voice Pathology Detection and Multi-classification Using Machine Learning Classifiers

Wu Yuanbo
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
ybwu0513@stu.suda.edu.cn

Zhou Changwei
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
20194239030@stu.suda.edu.cn

Fan Ziqi
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
zqfan@stu.suda.edu.cn

Zhang Yihua
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
yhzhang1001@stu.suda.edu.cn

Zhang Xiaojun†
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
zhangxj@suda.edu.cn

Tao Zhi*
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
taoz@suda.edu.cn

Abstract—**Automatic voice pathology detection can provide objective estimation and prevention in the early stages of voice diseases. A voice pathology detection and multi-classification method using an audio feature set extracted from glottal flow waveform is proposed in this work. In addition, a feature selection method of Fisher discrimination criterion is used to screen more valuable features and eliminate the features with low efficiency. The discrimination ability and effectiveness of the selected features were verified using three dimensional scatter plot and box plot in this study. One of the contributions of this paper is to investigate and evaluate the performance of different machine learning classifiers value for voice pathology detection and multi-classification. All experiments were carried out using the Massachusetts eye and ear infirmary database. Accuracy, sensitivity, specificity and receiver operating characteristic area are used as evaluation indexes to compare the performance of different machine learning classifiers. Each machine learning classifier had a good performance using the proposed feature set with nearly 100% accuracy in detection and higher than 90% accuracy in multi-classification. The experimental results indicate that the proposed method is valuable for voice pathology detection and multi-classification, especially in multi-classification.**

Keywords—*Voice pathology detection and classification, glottal flow waveform, machine learning methods, feature selection.*

I. INTRODUCTION

According to statistics, about 6-15% of the general population have a voice disorder because of misuse of pronunciation, improper use of voice, local infection and mental problems [1]. For certain professions, the proportion increase to even 50%, such as teachers, singers and announcers [2]. Compared with medical examinations, the method of voice pathology detection and classification using signal processing and machine learning is non-invasive and objective, which is the main focus of this paper.

Pathological voice detection and classification is generally conducted by analyzing and learning the acoustic characteristics of speech signals and using supervised machine learning to classify speech signals. Eskidere et al. used multitaper Mel Frequency Cepstral Coefficients (MFCCs) to detect voice diseases. The method was applied

on Saarbruecken Voice database (SVD), and the highest accuracy of 99.38% for vowel /a/ was achieved using Gaussian Mixture Model (GMM) and various windowing methods[3]. In [4], an accurate and robust method of feature extraction for detecting and classifying pathological voice by using autocorrelation and entropy in different frequency bands was proposed. Three different database, Massachusetts eye and ear infirmary (MEEI), SVD and Arabic voice pathology database (AVPD) were used and the highest accuracies 99.69%, 92.69% and 99.79% were achieved using support vector machine (SVM). In [5], vocal tract area directly connected to vocal folds was used as input parameters of the classifier to detect pathology. The method was evaluated using the MEEI database and SVD database, and achieved 99.22% and 94.7% accuracy based on SVM classifier, respectively. Jothilakshmi proposed an automatic voice pathology detection system using GMM and HMM. MFCC and linear prediction cepstral coefficients are used as input characteristics. For binary detection, this method can achieve the highest 94.44% efficiency, while the performance of classification between different types of pathology is giving highest sensitivity of 81.81% and specificity of 97.59% [6]. Obviously, the key factors influencing this classification method are whether the features used can fully reflect the difference between pathological and normal voices and the learning ability of the classifier used.

Cepstral features like MFCC and nonlinear features like autocorrelation were widely used and improved in previous work, and satisfactory results have been achieved with machine learning classifiers such as SVM and GMM. However, it is necessary to point that most researches focused on the general classification of normal and pathological voices composed of various types of voice diseases. The reason why spectral and cepstral features have no significant difference in the classification of different voice diseases may be that it focuses on the characteristics of vocal tract resonance [7]. Therefore, the features extracted from the glottal source signal have been studied in the detection and classification of voice pathology. Forero et al. used glottal signal parameters as input to an Artificial Neural

---
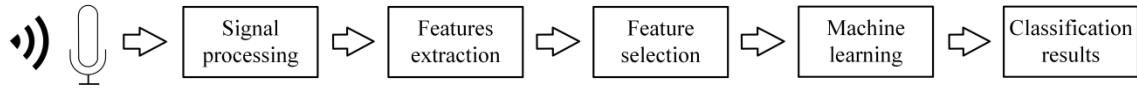
Fig. 1   The block diagram of the proposed method.

Network (ANN), an SVM and a HMM to analyze and classify voice pathologies. The database used here was composed of 248 voice recordings and the highest accuracy of 97.2% was achieved[8]. In [9], a method based on interlaced derivative pattern (IDP) of the glottal source was proposed to detect and classify voice pathology. The SVM with radial basis function (RBF) as a kernel was used. Compared with MFCC and MDVP, IDP achieved higher 88.5% accuracy for detection and 90.3% for classification. We can find that the features extracted from the glottal source signal can achieve good performance on both voice pathology detection and binary classification. In addition, research on pathological voice detection is relatively mature, while research on pathological voice multi-classification is relatively few.

As a consequence, the main contribution of this paper is to further study how to effectively multi-classify different voice diseases with glottal source features based on machine learning methods. Glottal inverse filtering and MPEG-7 audio standard were applied to extract the feature set used in this work. In addition, a feature selection method of Fisher discrimination criterion (FDC) was used to optimize the feature set. Another contribution of this paper is to investigate and evaluate the performance of different machine learning classifiers in voice pathology detection, especially multi-classification. The rest of this article is organized as follows. In section II, the methods of features extraction are introduced, and three machine learning classifiers employed in the experiments are described. Section III describes the feature selection method used and discusses the results compared with previous work. Finally, the work is concluded in Conclusion section V.

II.   METHODS AND MATERIALS

This section describes the proposed method and materials used in this paper. Fig. 1 shows the block diagram of the proposed method. The voice signal will be firstly processed by resampling and inverse filtering, then the extraction of features is conducted by time and frequency measuring and low-level descriptor extractor. Additionally, a feature selection method of FDC value ranking is used to optimize the feature set extracted. The selected features are then used as input parameters of machine learning classifiers and the experimental results will be discussed in the subsequent section.

A. Database

The main objective of this work is to further study the performance of the features extracted from the glottal source signal in voice pathology detection and multi-classification with machine learning methods. Therefore, in order to provide reference and comparison for future research, the database selected for the experiments should be widely used in the field and be generally recognized. The MEEI database[10] used in this paper has been widely applied in the work of pathological voice detection and classification. The MEEI database consists of sustained vowel /a/ by 53

normal speech recordings which sustain approximately 3 seconds and 657 pathological speech recordings which last around 1 second.

In most of the previous work, the databases used for classification of voice diseases were usually composed of just two categories: speakers with normal voice and speakers with voice diseases which include various kinds of diseases. However, binary detection and multi-classification between normal voice and four different kinds of voice diseases (vocal nodules, vocal polyps, vocal cord paralysis and vocal cysts) are considered in this work. This subset of the MEEI database contains sustained vowel /a/ recordings for 53 normal, 19 nodular, 20 polypous, 67 paralyzed and 44 cystial recordings. A statistical description of the experimental subset of the MEEI database is shown in TABLE.I.

TABLE I.   STATISTICS OF THE SUBSET USED OF THE MEEI DATABASE.

Voice diseases	Number		Mean age (years)		Age range (years)	
	Female	*Male*	*Female*	*Male*	*Female*	*Male*
Normal	32	21	35.17	37.80	22-52	26-59
Nodule	18	1	29.63	47.00	18-48	47-47
Polyp	8	12	44.83	45.47	30-74	21-76
Paralysis	34	33	54.33	53.58	19-80	15-77
Cyst	33	11	41.27	40.16	17-85	23-63

B. Features Extraction

During the speech signal production, vocal folds oscillate in a frequency called fundamental frequency when the airflow passes through. The airflow changes into the air pulses because of the oscillation of vocal folds. The pressure signal formed by the air pulses is quasi-periodic and it is called the glottal flow waveform[11]. The features are composed of glottal source features and MPEG-7 audio features, which are extracted from the glottal flow waveform computed by the inverse filtering algorithm called PSIAIF [12] available as built-in function of Aparat software [13].

The glottal source features obtained by PSIAIF are divided into two groups: time-domain features and frequency-domain features[14]. The time-domain features can be described by using different phases (closed phase, opening phase and closing phase) and instants (closure time instant, primary opening instant, secondary opening instant and the instant of minimum and maximum glottal flow). In addition, the amplitude of the glottal source signal in different phases is also used to compute the amplitude-based glottal source features. Different from time-domain features, frequency-domain features which reflect the slope of the spectrum in essence are calculated from the spectrum of the glottal flow. The detailed calculation can be found in [14].

MPEG-7 low-level audio features are derived from an international multimedia content description standard ISO/IEC 15938, which is composed of video and audio parts [15]. MPEG-7 standard has been applied on various acoustic research [16—18]. Inspired by the satisfactory performance of features obtained through MPEG-7 standard in above literatures, we applied this standard in glottal source flow to extract the audio features. TU-Berlin MPEG-7 audio

analyzer [19] was used to extract MPEG-7 features from glottal flows. The MPEG-7 low-level audio features consist of 17 descriptors which can be divided into two different categories: scalar and vector. A detailed and complete description of MPEG-7 low-level audio features can be found in [16]. The list of all 57 audio features extracted from glottal flow is shown in TABLE II.

TABLE II. THE NAME, TYPE AND SIZE OF THE FEATURES EXTRACTED FROM THE GLOTTAL SOURCE SIGNAL.

Name	Type&Size	Name	Type&Size
Primary Open quotient (OQ1)	Scalar	Audio Spectrum Centroid (ASC)	Scalar
Secondary Open quotient (OQ2)	Scalar	Audio Spectrum Spread (ASS)	Scalar
Normalized amplitude quotient (NAQ)	Scalar	Audio Harmonicity (AH)	Scalar
Amplitude quotient (AQ)	Scalar	Audio Fundamental Frequency (AFF)	Scalar
Closing quotient (ClQ)	Scalar	Temporal Centroid (TC)	Scalar
Open quotient (OQa)	Scalar	Spectral Centroid (SC)	Scalar
Quasi-open quotient (QoQ)	Scalar	Harmonic Spectral Centroid (HSC)	Scalar
Primary Speed quotient (SQ1)	Scalar	Harmonic Spectral Deviation (HSD)	Scalar
Secondary Speed quotient (SQ2)	Scalar	Harmonic Spectral Spread (HSS)	Scalar
Amplitude difference between the first two harmonics (H1-H2)	Scalar	Harmonic Spectral Variation (HSV)	Scalar
Parabolic spectral parameter (PSP)	Scalar	Audio Spectrum Flatness (ASF)	Vector & 22 features
Harmonic richness factor (HRF)	Scalar	Audio Spectrum Basis (ASB)	Vector & 2 features
Minimum of Audio Waveform (AW.min)	Scalar	Audio Spectrum Projection (ASP)	Vector & 2 features
Maximum of Audio Waveform (AW.max)	Scalar	Audio Spectrum Envelope (ASE)	Vector & 3 features
Audio Power (AP)	Scalar		

C. Machine Learning Classifiers

In order to investigate and evaluate the effect of different machine learning methods in the detection and multi-classification of pathological voices, SVM, Decision Tree (DT) and Random Forest (RF) were chosen to be used in this work. Their principles and methods are described below.

SVM is a dichotomous model which is widely used for data classification. Its basic model is the linear classifier with the largest interval defined in the feature space. SVM also includes kernel functions, which make it essentially a nonlinear classifier. The learning algorithm of SVM is the optimization algorithm for solving convex quadratic programming. The kernel of SVM used in this paper is radial basis function (RBF) and optimum parameters values for C = 2 and $\gamma = 0.1$ are computed by grid search.

Decision tree is a tree structure in which each internal node represents a judgment on an attribute, each branch represents the output of a judgment result, and finally each leaf node represents a classification result. The generation algorithm we used is J48, which is the most popular tree classifier.

The basic principle of RF classifier is to combine a certain amount of binary decision trees and output the results of classification by a majority voting [20]. It is important to point out that different bootstrap samples of the original data and a random subset of feature set are used to build the trees. If an iteration threshold is not set, the trees will continuously process until the elements in the leaves are from the same category. The number of trees in RF is denoted as k and the maximum number of randomly selected features is denoted as q. In this work, $k = 1000$ trees are used to build an RF and q is set to $q = \sqrt{p}$, which was recommended in [21].

For all machine learning classifiers, 10-fold cross-validation is used to carry out the experiments in this work. We divided the data set composed of many samples into 10 pieces of equal size and selected 9 pieces for training set and 1 piece for test set for each time. A total of 10 experiments were conducted. Finally, the mean value of the results of 10 experiments was taken as the final accuracy of classification.

III. RESULTS AND DISCUSSION

A. Feature Selection

If the number of all features in a feature set or space is relatively large, it is obvious that a part of features perform better whereas the other may perform worse and the time cost may dramatically increase if all features are considered. Considering the adverse effects of redundant features, feature selection is an effective method to select the more suitable features and reduce the dimension of the feature set or space to improve the relevance of the features at the same time. Hence, a feature selection method in the form of FDC is used to select more effective features. FDC for each feature which is denoted by k can be computed as follow.

$$\text{FDC}(k) = \frac{(\mu_{n,k} - \mu_{p,k})^2}{\sigma_{n,k} + \sigma_{p,k}} \tag{1}$$

where $\mu_{n,k}$ and $\mu_{p,k}$ express the mean values for classes of normal voice and pathological voice respectively, while $\sigma_{n,k}$ and $\sigma_{p,k}$ represent the variances for classes normal and pathological, respectively. If a certain feature can achieve a satisfactory performance in classification between two classes, the distance between the means of them should be high and the variance within each class should be low. That means a high value of FDC should be obtained. The rank of FDC values of the audio features extracted is shown in Fig. 2.

It can be seen that ASF10 (the 10th band of ASF), ASF9 (the 9th band of ASF) and ASC have the highest discrimination between normal and pathological voice. The number of features having FDC > 0.7 is 8 and all of them belong to MPEG-7 features. Most glottal source features can achieve good performance which obatin a minimum FDC value of 0.37 (PSP) and a maximum FDC value of 0.68 (AQ). AQ, OQ1 and ClQ, which are related to the phonation and voice intensity of speakers, are the top 3 glottal source features obtaining FDC of 0.68, 0.63 and 0.59, respectively. It should be pointed out that the total number of glottal source features is only 12 (about 1/4 of the total number of MPEG-7 features) and 9 of these features rank in the top 28
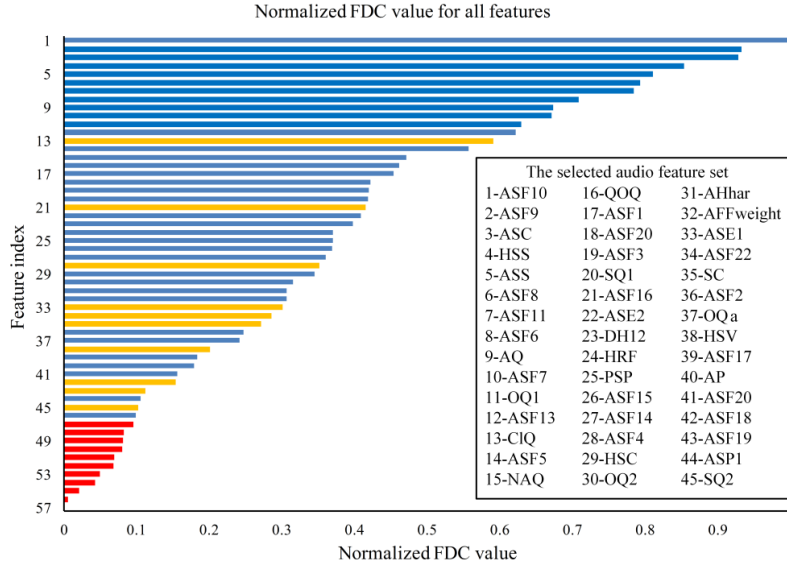
Fig. 2 Normalized FDC values of all the 57 audio features extracted from glottal flow waveform. Blue bars refer to the selected MPEG-7 audio features, Orange bars refer to the selected glottal source features, Red bars refer to features that are removed with FDC < 0.1.

(about half of the total 57 features) which indicates that most glottal source features can contribute to pathology detection and classification.

We removed the bottom 12 features with FDC < 0.1 which have little contribution for pathology classification and the rest 45 features constitute the selected audio feature set.

B. Pathology Detection Using Machine Learning Classifiers

Features extracted from glottal source signal have been used as input parameters for three machine learning classifiers: SVM, DT (J48) and RF. Results of voice pathology detection are shown in TABLE III. ACC, SE, SP and AUC which are considered as classical indexes have been applied to evaluate the performance of machine learning classifiers. Obviously, all the machine learning methods achieved good performance on the detection of pathological voice with the top 45 features selected by FDC ranking. Compared with the other two classifiers, RF have relatively high ACC with a difference of 0.71% and 0.83%, respectively. The SE, SP and AUC of each machine learning classifier are close to 1.00, which indicates that the machine learning methods used can effectively discriminate normal and pathological voices.

TABLE III. RESULTS OF PATHOLOGY DETECTION USING MACHINE LEARNING CLASSIFIERS WITH THE SELECTED FEATURES EXTRACTED FROM THE GLOTTAL SOURCE SIGNAL. SE REFERS TO SENSITIVITY, SP REFERS TO SPECIFICITY AND AUC REFERS TO AREA UNDER THE RECEIVER OPERATING CHARACTERISTIC CURVE.

Machine learning Classifier	ACC(%)	SE	SP	AUC
SVM	99.28	0.99	1.00	1.00
DT (J48)	99.16	0.99	1.00	1.00
RF	**99.99**	**1.00**	**1.00**	**1.00**

Fig. 3 describe a 3D scatter plot of the values for top 3 features ranked by FDC (ASC, ASF9, ASF10). We can find that the distribution of pathological voices represented by red dots is clearly separated from the distribution of normal voices represented by blue dots. This phenomenon can greatly prove that the selected features extracted from the

glottal source signal have a satisfactory discrimination ability between normal and pathological voices.
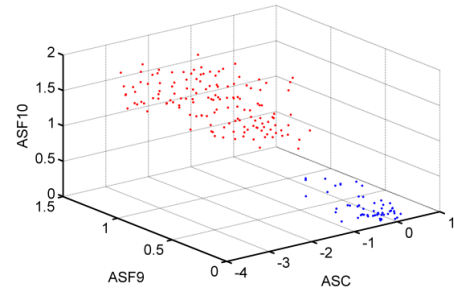
Fig. 3 Three dimensional scatter plot of ASC, ASF9 and ASF10 features for each voice sample in the database used. Red dots refer to the pathological voices and blue dots refer to the normal voices.

C. Pathology Binary Classification

To evaluate the effectiveness of the selected feature set in the binary classification of 4 different voice diseases, 6 binary experiments have been constructed with different machine learning methods as shown in TABLE IV. As we can observe, all three machine learning methods achieved good performance in binary classification of voice pathology using the proposed feature set. The average accuracy of 97.81% obtained by Random Forest classifier is the highest accuracy compared with 95.93% of SVM and 97.10% of DT (J48). The RF method had the best performance in most binary classification experiments except two ones. Compared to 95.97% average accuracy of the methods proposed in [9], the classification accuracy for the same diseases reaches 98.86% using the MEEI database.

It is worth noting that vocal paralysis is more distinct from the other three diseases. The reason for this phenomenon may be that vocal folds paralysis is a psychiatric voice disease, which is fundamentally different in pathological mechanism from the hyperplastic voice disease. These differences directly affect the acoustic characteristics of paralytic and non-paralytic voice signals, and make it easier to distinguish between the two types of samples, which is consistent with the experimental results in [18].

Diseases categories	Classifiers	ACC(%)	SE	SP
Nodule vs. Polyp	SVM	94.28	0.94	0.96
	DT(J48)	**95.28**	**0.95**	**0.96**
	RF	94.95	0.95	0.95
Nodule vs. Paralysis	SVM	99.18	0.97	0.99
	DT(J48)	99.27	0.98	0.99
	RF	**99.82**	**0.99**	**1.00**
Nodule vs. Cyst	SVM	91.95	0.87	0.94
	DT(J48)	93.45	0.92	0.94
	RF	**95.54**	**0.93**	**0.96**
Polyp vs. Paralysis	SVM	98.33	0.93	0.99
	DT(J48)	98.48	0.96	0.99
	RF	**99.85**	**0.99**	**1.00**
Polyp vs. Cyst	SVM	96.00	0.91	0.98
	DT(J48)	**97.33**	**0.95**	**0.97**
	RF	96.92	0.95	0.98
Paralysis vs. Cyst	SVM	95.87	0.99	0.90
	DT(J48)	98.82	0.99	0.98
	RF	**99.80**	**1.00**	**0.99**

D. Pathology Multi-classification Using Machine Learning Classifiers

The main objective of this work is to explore the performance on pathology multi-classification by using different machine learning methods and evaluate the discrimination of the proposed audio feature set extracted from glottal flow waveform. Result for multi-classification of pathological voices using the features selected by FDC ranking is shown in TABLE V. As we can see that each machine learning method used in voice pathology multi-classification achieved an accuracy higher than 90%. The highest ACC of 93.89% has been achieved by the RF classifier and the lowest ACC of 90.82% has been achieved by the SVM classifier. In addition, the accuracy obtained by DT (J48) is close to the one in RF and higher than SVM. This indicates that the performance of tree classifier (DT) is better than that of linear classifier (SVM), and the integrated learning classifier (RF) has further excellent performance. The increase of accuracies of three machine learning methods may because that the high detection rate of normal speech raises the overall accuracy, when normal voice samples were added in.

Fig. 3 shows the distributions of the top 8 audio features ranked by FDC value: ASF10, ASF9, ASC, HSS, ASS, ASF8, ASF11 and ASF6. It can be seen that all the top 8 features achieve really good performance and high discriminations between normal and pathological voice (Nodule, Polyp, Paralysis and Cyst). Apart from ASC, the distribution range of the each remaining feature in normal voice is generally less than the one in pathological voice, and the bottom whisker of pathology almost exceeds the top whisker of normal voice. For ASF10, ASF9 and ASF11, although the difference between the maximum and minimum of each box is relatively close to each other, the median value of each box which represents the general evaluation level of each feature to some extent can be easily distinguished from each other. ASC shows that the range of feature value in vocal nodule is obviously lower than the others, and the features value in the other 3 diseases has larger variations compared to the one in nodule. For HSS feature, all 4 types show relatively lower variations, but the distributions of them are different from each other since

nodule and paralysis center on the upper part while polyp and cyst are concentrated in the lower part. ASS, ASF8 and ASF6 show large differences in distributions between nodule, polyp, paralysis and cyst, but there are some outliers in polyp and cyst for ASS and in nodule and cyst for ASF8.

Diseases categories	Machine learning classifiers	ACC(%)	SE	SP
Nodule, Polyp, Paralysis and Cyst	SVM	90.82	0.86	0.96
	DT (J48)	93.35	0.90	0.96
	RF	**93.89**	**0.91**	**0.97**
Normal, Nodule, Polyp, Paralysis and Cyst	SVM	91.45	0.87	0.96
	DT (J48)	96.75	0.91	0.98
	RF	**97.11**	**0.93**	**0.99**

To investigate the effectiveness of machine learning classifier in the multi-classification of pathological voices, the confusion matrix of the RF classifier which had the best performance has been calculated shown in TABLE VI. The leading diagonal of the confusion matrix expresses the percentage of correctly classified samples of 4 different types of voice diseases. As we can see that the numbers of correctly classified samples of vocal fold paralysis is most and only few cyst samples were confused with paralysis. However, nodule, polyp and cyst were found to be confused between each other, especially nodule and cyst. We suspect that in addition to the similarity of symptoms, the imbalance in the number of pathological samples involved in the training of classifier also affected the results of the confusion matrix.

In summary, the proposed audio feature set selected by FDC value ranking is valuable for voice pathology research. All machine learning classifiers used in this work achieved satisfactory performance. Overall, the RF classifier is more effective than SVM and DT for both pathology detection and multi-classification.

Output→ Input↓	Nodule	Polyp	Paralysis	Cyst
Nodule	**79.79**	5.00	0.37	14.84
Polyp	4.80	**90.10**	0.30	4.80
Paralysis	0	0	**99.98**	0.02
Cyst	6.78	2.29	0.33	**90.60**

IV. CONCLUSION

A method of detection and multi-classification of voice pathology based on the audio feature set extracted from glottal flow waveform was proposed in this paper. The proposed method achieved 99.99%, average 97.81% and 93.89% accuracy for detection, binary classification and multi-classification, respectively. The strong discrimination ability of the features selected by FDC value ranking was proved by analysis of the 3D scatter plot and box plot. Additionally, this paper investigates and evaluates the performance of three machine learning classifiers: SVM, DT and RF. The experimental results show that the evaluated three machine learning methods can effectively distinguish different types of pathological voices, and the performance of the tree classifier is better than that of the linear classifier. The high accuracy for multi-classification between various

voice diseases signifies that the proposed method can play the role of a medical tool for helping doctors to diagnose disorders well.
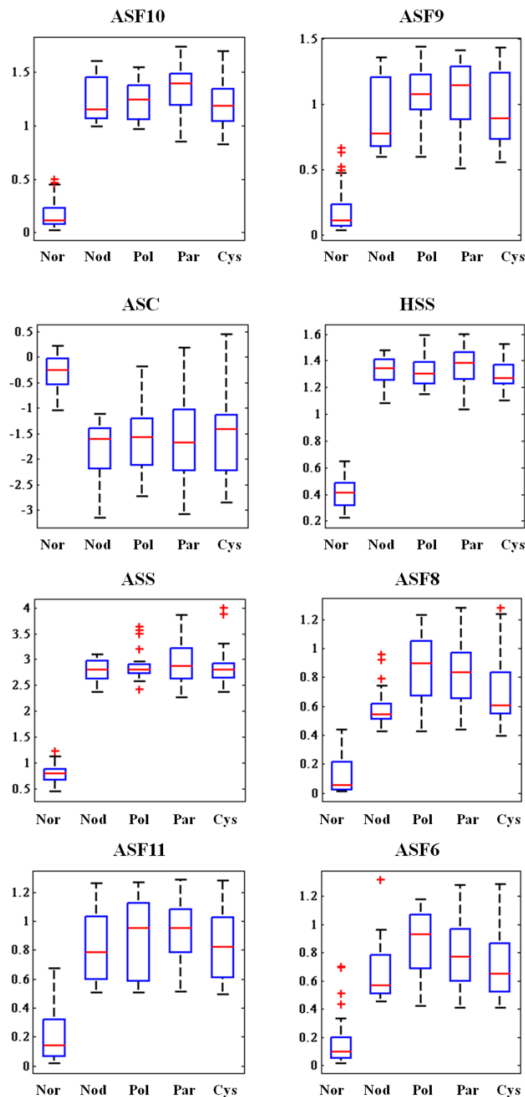
Fig. 4 Distribution of the top 8 features selected from the audio feature set using box plots. Nor, Nod, Pol, Par, Cys represent normal voice, vocal nodule, vocal polyp, vocal folds paralysis and vocal cyst, respectively. The central read line of the box represents the median, and the bottom and top edges of the box represents the 25% and 75% values, respectively. All the sample points within 3/2 times the interquartile range are covered by the whiskers on the top and bottom sides. The sample points beyond these whiskers are marked as outliers using the '+' symbol.

Considering the discrimination of the proposed method for voice pathology detection and classification, it is worth to investigate its performance in detecting speech disorders, Parkinson's dysphonia and even laryngeal canceration. A future work will be to study the imbalance of numbers of voice pathology, which leads to the difficulties of classification. In addition, considering the potential limitation of the MEEI database such as language, recording environment and setups, more databases and more categories of voice pathologies will be used in the future.

REFERENCES

[1] Garcia Martins, R. H., Boia Neves Pereira, E. R., Hidalgo, C. B., et al. Voice Disorders in Teachers. A Review[J]. Journal of Voice, 2014, 28(6).

[2] Cutiva, L. C. C., Vogel, I., Burdorf, A. Voice disorders in teachers and their associations with work-related factors: A systematic review[J]. Journal of Communication Disorders, 2013, 46(2): 143-155.

[3] Eskidere, O., Gurhanli, A. Voice Disorder Classification Based on Multitaper Mel Frequency Cepstral Coefficients Features[J]. Computational and Mathematical Methods in Medicine, 2015.

[4] Al-Nasheri, A., Muhammad, G., Alsulaiman, M., et al. Voice Pathology Detection and Classification Using Auto-Correlation and Entropy Features in Different Frequency Regions[J]. Ieee Access, 2018, 6: 6961-6974.

[5] Muhammad, G., Altuwaijri, G., Alsulaiman, M., et al. Automatic voice pathology detection and classification using vocal tract area irregularity[J]. Biocybernetics and Biomedical Engineering, 2016, 36(2): 309-317.

[6] Jothilakshmi, S. Automatic system to detect the type of voice pathology[J]. Applied Soft Computing, 2014, 21: 244-249.

[7] Kadiri, S. R., Alku, P. Analysis and Detection of Pathological Voice Using Glottal Source Features[J]. Ieee Journal of Selected Topics in Signal Processing, 2020, 14(2): 367-379.

[8] Forero, L. a. M., Kohler, M., Vellasco, M. M. B. R., et al. Analysis and Classification of Voice Pathologies Using Glottal Signal Parameters[J]. Journal of Voice, 2016, 30(5): 549-556.

[9] Muhammad, G., Alsulaiman, M., Ali, Z., et al. Voice pathology detection using interlaced derivative pattern on glottal source excitation[J]. Biomedical Signal Processing and Control, 2017, 31: 156-164.

[10] Massachusetts Eye and Ear Infirmary, Elemetrics Disordered Voice Database (Version 1.03), Voice and Speech Laboratory, Boston, MA, 1994, Available at http://www.kayelemetrics.com/.

[11] Rosa Mde O, Pereira JC, Grellet M. Adaptive estimation of residue signal for voice pathology diagnosis. IEEE Trans Biomed Eng. 2000;47:96–104.

[12] Raitio T, Suni A. Hmm-based speech synthesis utilizing glottal inverse filtering. IEEE Trans Audio Speech Lang Processing.2010; 19:153–165.

[13] Software Aparat, Available at: http://aparat.sourceforge.net/index.php/Main_Page, Helsinki University of Technology Laboratory of Acoustics and Audio Signal Processing, Accessed June 15, 2015.

[14] Airas, M. TKK Aparat: An environment for voice inverse filtering and parameterization[J]. Logopedics Phoniatrics Vocology, 2008, 33(1): 49-64.

[15] Information Technology-Multimedia Content Description Interface-Part 4:Audio, ISO/IEC CD 15938-4, 2001.

[16] Sarno, R., Ridoean, J. A., Sunaryono, D., et al. Classification of Music Mood Using MPEG-7 Audio Features and SVM with Confidence Interval[J]. International Journal on Artificial Intelligence Tools, 2018, 27(5).

[17] Luque, J., Larios, D. F., Personal, E., et al. Evaluation of MPEG-7-Based Audio Descriptors for Animal Voice Recognition over Wireless Acoustic Sensor Networks[J]. Sensors, 2016, 16(5).

[18] Muhammad, G., Melhem, M. Pathological voice detection and binary classification using MPEG-7 audio features[J]. Biomedical Signal Processing and Control, 2014, 11: 1-9.

[19] TU-Berlin MPEG-7 Audio Analyzer, Technische Universität Berlin, 2014, Available at http://mpeg7lld.nue.tu-berlin.de/.

[20] Vaiciukynas, E., Verikas, A., Gelzinis, A., et al. Fusion of voice signal information for detection of mild laryngeal pathology[J]. Applied Soft Computing, 2014, 18: 91-103.

[21] Breiman L. Random forests. Machine Learn. 2001;45:5–32. 10.1023/A:1010933404324.

Classification of Normal and Pathological Voices Using Convolutional Neural Network

Zhou Changwei
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
20194239030@stu.suda.edu.cn

Zhang Lili
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
20184208133@stu.suda.edu.cn

Zhang Xiaojun
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
zhangxj@suda.edu.cn

Wu Yuanbo
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
ybwu0513@stu.suda.edu.cn

Wu Di†
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
wudi@suda.edu.cn

Tao Zhi*
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
taoz@suda.edu.cn

Abstract—In this paper, the spectrogram features of voice are studied, and the differences between normal voice and pathological voice are analyzed. 53 cases of normal voice and 133 cases of pathological voice are selected from Massachusetts Eye and Ear Infirmary(MEEI) database for experiments. As the popularity of deep learning is increasing, convolutional neural network(CNN) based on tensorflow2.1 is applied to extract features from spectrograms., and the voice data are augmented for overfitting of the trained model. The recognition accuracy of pathological voice detection is 96.51%. The experimental results show that the characteristic parameters extracted from spectrograms by CNN have distinguish ability.

Keywords—pathological voice, spectrogram, convolutional neural network, feature extraction

I. Introduction

Voice is the most natural, effective way for communication between human beings. It is particularly important for the occupations who rely on the voice for their livelihood, such as teachers, announcers, singers, actors etc. However, with the disturbance of living habits and overuse of vocal folds, the incidence of voice disorders is rising, which affects voice quality and phonation[1]. Thus, more importance should be attached to early diagnosis of vocal fold diseases. Although there are clinical diagnostic methods, such as stroboscopic laryngoscope, laryngeal electromyography and high-speed photography, these methods are invasive and painful to patients, so a painless and automatic method plays an importance role on the early detection of vocal fold diseases.

The healthy vocal folds vibrate periodically, and the vocal folds with some diseases have the irregularities of vibration[2], which results in the periodical change and amplitude change of vocal cord vibration. To detect voice pathologies automatically, researchers have used several voice features. As the conventional acoustic parameters, Jitter describes the instability of the vibration frequency of the vocal folds. Shimmer is similar to Jitter, which represents the variation of amplitude between adjacent vibration cycles. They are associated with roughness and hoarseness of voice. Due to the irregularity of the pathological voice, harmonic to noise ratio (HNR) is also an effective feature, which shows the relationship of voice between periodic and non-periodic

signals. Recently, the other common voice features to detect voice diseases are Mel-Frequency cepstral coefficients (MFCC)[3-5], linear prediction cepstral coefficients (LPCC)[6] and so on. They do well in characterize the shape of the vocal tract. However, most of the acoustic features extracted based on signal processing algorithms are artificially defined, which are subjective and require a lot of time and energy for verification. On the other hand, the inherent characteristics of the voice cannot be well represented. As a result, the extracted features are not complete enough to reflect the characteristics of the voice, which lead to a relatively low recognition accuracy.

In this paper, the spectrogram is analyzed to detect normal and pathological voices, which includes a sequence of different features that vary with time. Convolutional Neural Network(CNN) is applied for extracting features of the spectrogram, which can be completed automatically through different convolutional layers. Compared with shallow neural networks such as back-propagation neural networks, CNN is able to extract complex and abstract features, and can avoid falling into local optimal solution model[7].

The rest of the paper is organized as follows. Section II analyzes the differences between spectrograms of normal and pathological voices in detail. In section III, the database and the structure of CNN are presented. Section IV discusses the results. Finally, conclusions are drawn in section V.

II. Spectrograms of Normal and Pathological Voices

Human voice can be considered to be stationary signal in the period of 10-30ms. a sequence of different events varying with time are including in the speech waveform, which are associated with fluctuating spectral characteristics. Hence, the short time Fourier transform(STFT) is used for analysis[8].STFT divides the voice signal into segments of the waveform under a sliding window, which are processes by a single Fourier transform respectively.

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{+\infty} [x(m)\omega(n-m)]e^{-j\omega m} \tag{1}$$

where $x(m)$ is input voice signal at time m and $\omega(n-m)$ is the sliding window.

*Tao Zhi is the first corresponding author. (e-mail: taoz@suda.edu.cn)
†Wu Di is the second corresponding author. (e-mail: wudi@suda.edu.cn)
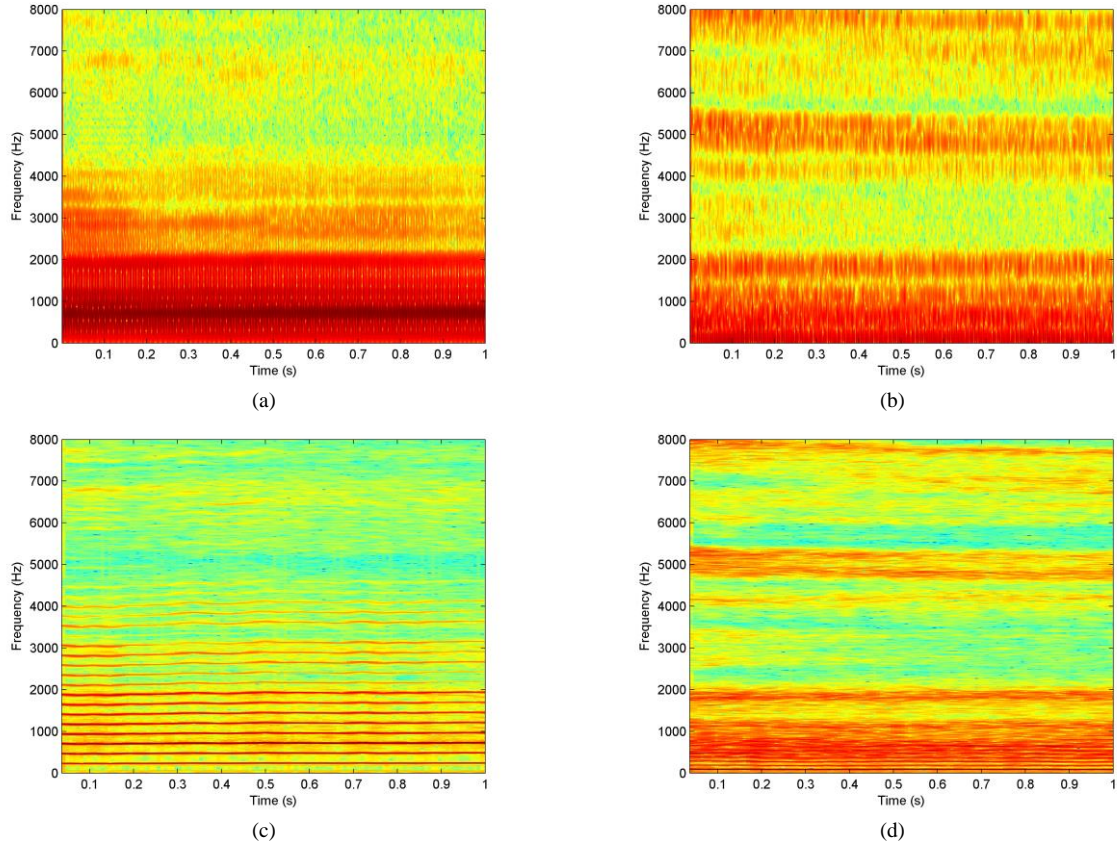
Fig. 1 Broadband spectrograms of a (a) normal and (b) pathological (polyp) voice of the vowel sound /a:/. Narrowband spectrograms of the same utterance (c) normal and (d) pathological voice.

Then, based on definition of power spectrum, the spectrogram S_n of the voice signal is derived from STFT by

$$S_n = \left| X_n(e^{j\omega}) \right|^2 \quad (2)$$

The broadband and narrowband spectrograms of a normal and pathological voice are shown in Fig. 1, which present the distribution of power densities with time and frequency.

The narrowband spectrogram can display the structure of harmonics clearly, and reflect the variety of the fundamental frequency with time, while the broadband spectrogram displays the formant and the envelope of spectrum. Comparing Figure 1(a) with Figure 1(b), the energy density of voice has been changed in that vocal polyps destroy the regularity of voice pronunciation, and change the amplitude range of voice. It can also be seen from the spectrogram in Figure 1(c) and Figure. 1(d) that the vocal polyps introduce high frequency noise and the energy diffuses from low frequency to high frequency. and produce high-frequency oscillation similar to noise.

Spectrograms of normal and pathological voice are different in energy density, frequency and time domain, which can be manifested from texture features of spectrograms. The grey level co-occurrence matrix(GLCM)[9] of image can reflect the comprehensive information of the gray-level of the image about the direction, the adjacent interval, and the range of variation. The grey level co-occurrence matrix is the descriptor for analyzing the local texture and arrangement rules of images. There are some texture features based on normalized GLCM, such as contrast, correlation, energy and homogeneity. Contrast reflects the depth of texture grooves and the sharpness between gray level of pixels. Energy is the sum of the squares of the pixels in GLCM, reflecting the uniformity of the gray level distribution of the image. Homogeneity measures the local uniformity of the image. If the gray level of each pair of pixels is similar, the uniformity value is large.

TABLE I.   TEXTURE FEATURES OF NARROWBAND SPECTROGRAMS

Category	Contrast	Correlation	Energy	Homogeneity
Normal	0.093	0.906	0.524	0.963
Pathological	0.265	0.784	0.252	0.886

TABLE. I shows the different texture features of spectrograms of normal and pathological voice, which are extracted from Fig. 1(c) and Fig. 1(d). The contrast and energy of spectrograms of normal and pathological voice are clearly different, which indicates that the texture of spectrogram in normal voice is more uniform than pathological voice. The differences between texture features in TABLE. 1 further demonstrate the feasibility of classification of normal and pathological voices by the spectrograms. The spectrograms will be a good indicator for the recognition of pathological voices. The more complex and abstract features of spectrograms will be extracted by convolutional neural network(CNN) for better recognition.
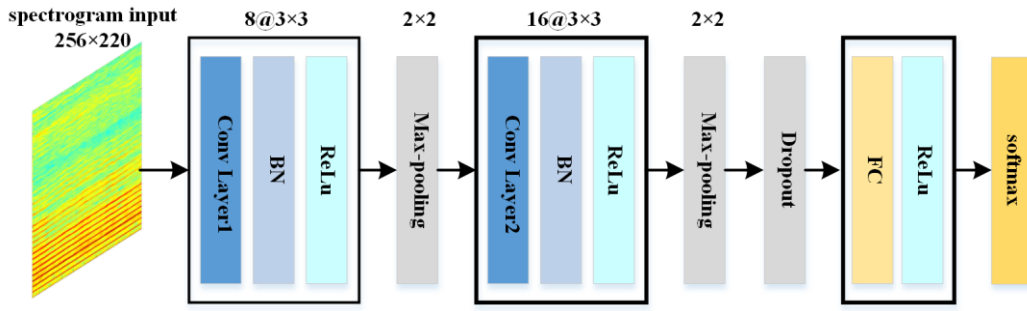
Fig. 2 Structure of normal and pathological voice classification by CNN

III. FEATURES EXTRACTION OF SPECTROGRAM BY CNN

A. Database

The voice data used in this paper are chose from MEEI database[10]. This database contains 1384 pathological voices, which gives the information of age, sex, smoking and other information of patients. The train and test set of normal and pathological voice are vowel /a:/. Three types of pathological voices are selected as pathological group, which consists of vocal fold polyps, vocal fold edema and paralysis. We used 53 cases of normal voices and 133 cases of pathological voice (21 vocal fold polyps, 44 vocal fold edema and 68 vocal fold paralysis). The sampling rate of normal voice is 50kHz, and the sampling rate of pathological voice is 25kHz or 50kHz with the bit width of 16 bits. For the consistency of voice, the sampling rate of all selected voices are dropped to 25kHz.

TABLE II.   STATISTICS OF VOICE DATA

Category		Numbers	Average age(STD)
Normal Voice		53	35.2(8.3)
Pathological Voice	Vocal Edema	44	45.6(10.9)
	Vocal Polyps	21	41.0(7.2)
	Vocal Paralysis	68	39.3(8.5)

The number of samples in the MEEI database is not enough for training an effective deep model. However, to obtain a large number of pathological voice is not practical, which involves the privacy of patients and needs the voice labeling by expert otolaryngologists. Hence, the data augmentation of voice samples is needed to increasing the number of the data. Ihsan Ullah augmented the data by splitting the full length electroencephalogram(EEG) signals into small slices using a fixed size window[11].Similarly, a window size of 5000 with a stride of 1000 is applied to the voice segmentation based on short-term stationary characteristics of human voices. The time of normal and pathological voice is 3s and 1~3s respectively. Finally, 756 cases of normal voice and 754 cases of pathological voice are obtained from the selected set in MEEI.

B. Architecture of CNN

Figure 2 shows the structure of normal and pathological voice classification by CNN. The input of CNN is the reshaped spectrograms with 256×220 pixels. CNN is established by convolutional layer, batch normalization(BN) layer, activation layer, pooling layer and dropout layer. Due to the number of voice samples, less convolution kernels and network layers were applied in CNN architecture for alleviating overfitting.

In this paper, the CNN architecture involves two convolution layers, two max-pooling layers and one dropout layer. The first convolutional layer uses 8 convolution kernels of size 3×3 to perform convolution operation on the reshaped spectrograms of size 256×220, and the stride is 1 pixel. A zero padding is set for ensuring the output size of spectrograms. The number of convolution kernels of the second convolutional layer increases to 16 for the smaller size of the feature map in deeper layer. The raise in number of convolution kernels expands the depth of the feature map, which maintains the capacity of carrying information.

As the number of network layer increase, features will deviate from zero mean. Hence, batch normalization is used after every convolution layer. BN makes the data conform to the distribution of 0 mean and 1 standard deviation.

$$H_i^{'k} = \frac{H_i^k - \mu_{batch}^k}{\sigma_{batch}^k} \tag{3}$$

where H_i^k is the pixel i of the feature map in the k convolution kernel before batch standardization, μ_{batch}^k, σ_{batch}^k are the average and standard deviation of all pixels in the batch of feature maps. In addition, scaling factor γ_k and offset factor β_k are added to optimize width and offset of data distribution, which solves the problem of gradient disappearance caused by equation (3). The fixed pixel X_i^k is as follows:

$$X_i^k = \gamma_k H_i^{'k} + \beta_k \tag{4}$$

In order to reduce parameter counts and the computational complexity, pooling layer is essential. In Fig. 2, the two pooling layers have pooling kernels of size 2×2, stride 2 pixel using maximization function. Finally, the feature map is formed into fully-connected layer and classification layer with softmax function. The function of all activation layers is rectified linear unit (ReLU).

IV. EXPERIMENTS AND RESULTS ANALYSIS

A. Experimental Setup

The training framework for classification of normal and pathological voices was established based on Tensorflow2.1, which is a very powerful platform and provides a variety of tools that makes it very easy to implement a model for machine learning applications in different areas. Training data was divided as 32 samples in each batch. The maximum epoch of training is 50. And an adaptive moment estimation (Adam)

optimizer was applied for good performance. In this experiment, 75% of the samples were selected as the training set to complete the training of normal and pathological voice classification models, and 25% of data were the test set.

The hardware configuration is as follows:

- Intel (R) Core(TM) i7-9750H @ 2.60GHz CPU.

- 16GB DDR4 memory.

- NVIDIA Geforce GTX1660Ti display card.

B. Classification of Normal and Pathological Voices by CNN

Figure 3 presents accuracy and loss function curve during the process of network training and test. After several iterations of the convolutional neural network, the accuracy on the training set is generally close to 100%, which indicates that the designed network can train well on MEEI databases and the data on the training set achieved a good fit. As the number of iterations increases, the accuracy of the training set and the test set is gradually rising. The value of the loss function is decreasing continuously, and eventually converges to a smaller value. And the smaller the value of the loss function is, the better the predicted value and the true value fit. When the normal and pathological voices are classified in the MEEI database, the accuracy of the validation set is 96.51%.
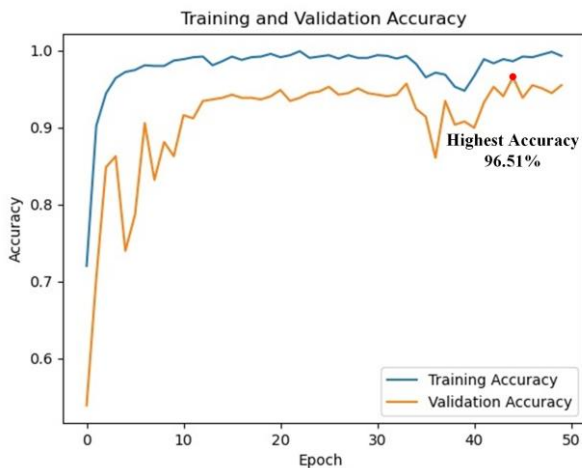
(a)
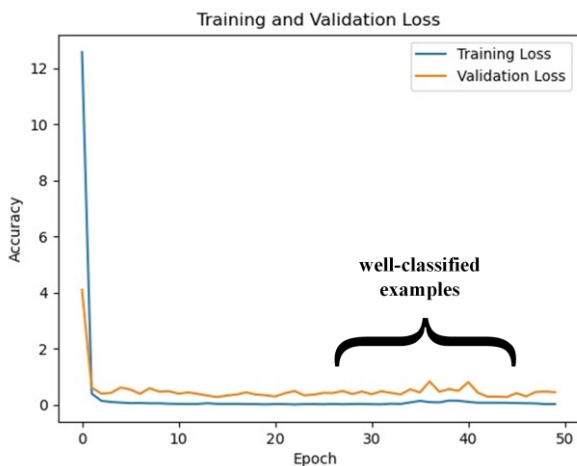
(b)

Fig. 3 The (a)accuracy and (b)loss curves of training and validation

C. Comparison with Other Methods

In order to better evaluate the result in this paper, the methods proposed in other literatures are compared with it in terms of extracted features, machine learning classifiers, and accuracy.

TABLE III.   COMPARISON OF DIFFERENT METHODS ON VOICE PATHOLOGY DETECTION

	Features	Classifier	Accuracy
Method [12]	Pitch, Jitter, Shimmer	Random Forest	90.27%
Method [13]	Modulation Spectral Features	SVM	94.1%
Method [14]	Mel-Spectrogram	CNN	93.8%
Our Method	Spectrogram	CNN	96.51%

TABLE III shows a comparison of our proposed method with three other methods[12-14] on the same subset of MEEI database. It can be clearly seen that the spectrogram feature outperforms the acoustic characteristics such as pitch, Jitter and Shimmer. And the recognition accuracy of the CNN used in this paper is higher than SVM classification, which reflects the advantages of CNN in feature extraction and classification. Classification of normal and pathological voices by CNN can avoid subjective factors, which enhances the robustness of features. However, due to the difficulties in obtaining medical samples, the number and categories of the samples are severely limited. When training deep models, small samples are prone to resulting in the over fitting of the model. Hence, we increased the sample size by data segmentation, bringing about better accuracy than method[14].

V. CONCLUSION

Classification of normal and pathological voices based on the spectrogram features by CNN are proposed in this paper. The proposed method achieved 96.51% accuracy for classification. The conventional features and classifiers are compared with our proposed method, which confirms the efficiency of extracting features from spectrogram by CNN. A deep learning algorithm is contributed to voice disability detection. And data augmentation is essential for classification of normal and pathological voices, which will improve accuracy and generalization of pathological voice detection.

In the future, the following aspects need to be further studied: (1) The differences of spectrograms between normal and pathological voice have been analyzed and achieved good performance. However, Jitter and Shimmer may need to be considered for pathological voice segmentation, which affect the speech spectrogram. (2) Small samples limit the training of deep learning model. transfer learning and more effective data augmentation could be explored.

REFERENCES

[1] Leborgne D L , Donahue E N . Voice Therapy as Primary Treatment of Vocal Fold Pathology[J]. Otolaryngologic Clinics of North America, 2019, 52(4):649-656.

[2]   Ulozaitestaniene N, Petrauskas T, Saferis V, Uloza V. Exploring the feasibility of the combination of acoustic voice quality index and glottal function index for voice pathology screening[J]. European Archives of Oto-rhino-laryngology, 2019, 276(6): 1737-1745.

[3]   Solanalavalle G, Galanhernandez J, Rosasromero R. Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features[J]. Biocybernetics and Biomedical Engineering, 2020, 40(1): 505-516.

[4]   Kadiri S R , Alku P . Analysis and Detection of Pathological Voice Using Glottal Source Features[J]. IEEE Journal of Selected Topics in Signal Processing, 2019: 1-13.

[5]   Verde L, De Pietro G, Sannino G. Voice Disorder Identification by Using Machine Learning Techniques[J]. IEEE Access, 2018: 16246-16255.

[6]   Hammami I, Salhi L, Labidi S. Pathological voices detection using Support Vector Machine[C]. international conference on advanced technologies for signal and image processing, 2016: 662-666.

[7]   Amin S U, Hossain M S, Muhammad G, Alhussein, M., Rahman, M. A. Cognitive Smart Healthcare for Pathology Detection and Monitoring[J]. IEEE Access, 2019: 10745-10753.

[8]   Kim B, Kong S, Kim S. Low Computational Enhancement of STFT-Based Parameter Estimation[J]. IEEE Journal of Selected Topics in Signal Processing, 2015, 9(8): 1610-1619.

[9]   Khaldi B, Aiadi O, Kherfi M L. Combining colour and grey-level co-occurrence matrix features: a comparative study[J]. Iet Image Processing, 2019, 13(9): 1401-1410.

[10]   Massachusetts Eye and Ear Infirmary, Elemetrics Disordered Voice Database(Version 1.03), Voice and Speech Laboratory, Boston, MA, 1994, Available at http://www.kayelemetrics.com/

[11]   Ullah I, Hussain M, Qazi E, Aboalsamh, H. An Automated System for Epilepsy Detection using EEG Brain Signals based on Deep Learning Approach[J]. Expert Systems With Applications, 2018: 61-71.

[12]   Zhang X, Zhang L, Tao Z, Zhao, H. Acoustic Characteristics of Normal and Pathological Voices Analysis and Recognition[C]. international conference on systems, 2019.

[13]   Maria Markaki, Yannis Stylianou. Voice Pathology Detection and Discrimination Based on Modulation Spectral Features[J]. IEEE Transactions on Audio Speech & Language Processing, 2011, 19(7):1938-1948.

[14]   Guan H, Lerch A. Learning Strategies for Voice Disorder Detection[C]. IEEE international conference semantic computing, 2019: 295-301.

Modeling Voice Pathology Detection Using Imbalanced Learning

Ziqi Fan
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
zqfan@stu.suda.edu.cn

Jinyang Qian
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
jyqian@stu.suda.edu.cn

Baoyin Sun
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
bysun@suda.edu.cn

Di Wu
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
wudi@suda.edu.cn

Yishen Xu*
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
xys2001@suda.edu.cn

Zhi Tao*
School of Optoelectronic Science and Engineering
Soochow University
Suzhou, China
taoz@suda.edu.cn

Abstract—To solve the problem of imbalanced data distribution of MEEI database, a modeling voice pathology detection method use imbalanced learning algorithm is presented in this paper. The method is based on support vector machine (SVM), decision tree (DT) and random forest (RF) to build voice pathology detection(VPD) model. Three imbalanced learning algorithms, SMOTE, Borderline-SMOTE and ADASYN, oversample a minority class (normal voice samples) and finally a training set of class-balanced is obtained. 10-fold cross validation is used in the experiment, and Precision, Recall and AUC/PRC are selected as model evaluation measures. Experimental results show that the imbalanced learning algorithm improves the recognition ability of pathological voice detection model for a minority class. RF as an ensemble model can combine better with imbalanced algorithms and performance slightly better than single classifier models. Its originality lies in the consideration of MEEI database as training set of pathological voice model, the effect of class-imbalanced on model performance, and Accuracy (*Acc*) as the main evaluation measure of VPD model is not suitable for training set of class-imbalanced

Keywords—Voice pathology detection model, Imbalanced database, Smart healthcare, Artificial intelligence, Imbalanced learning.

I. INTRODUCTION

In recent year, the automatic detection system based on machine learning has achieved noticeable performance in medical diagnosis. In the research of biomedical engineering, automatic voice pathology detection (VPD) system by machine learning algorithms and well-established features has become a research hotspot. Different types of signals are extracted for acoustic analysis [1, 2, 3] to realize automatic voice pathology detection.

Most of these studies are experimented with Massachusetts Eye and Ear Infirmary (MEEI) database [4]. Nonetheless, on MEEI database, there is class-imbalanced samples distribution of normal and pathological samples. Because of the diversity of pathological voice types, expensive endoscopy (such as laryngoscope) and labeling tasks with manual supervision, it is challenging to obtain a class-balanced data samples in practice. The learning of imbalanced dataset reduces the robust and generalization performance of classifiers, which is also one of the major challenges facing machine learning.

In the current study, there are three ways that are applied to solve this issue: improved algorithm [5, 6], cost sensitive learning [7] and data preprocessing [8, 9, 10]. Each algorithm has its unique strength, but compared with the other two methods, the data preprocessing method can perform better in the case of data instances. The synthetic minority over-sampling technique (SMOTE) [11] as one of the most classical data preprocessing algorithm, which is widely used in the preprocessing of imbalance datasets. In addition, Borderline-SMOTE [12], Adaptive synthetic sampling (ADASYN) [13] are extensions of SMOTE. In order to enhance the classification ability of VPD models for correct predictions, this study uses the oversampling's imbalance learning algorithms to process the MEEI database and select Mel-scale Frequency Cepstral Coefficients (MFCCs) as input feature parameters of model. The used of support vector machine (SVM) [14], decision tree (DT) [15] and random forest (RF) [16] classifiers to model MEEI training dataset using imbalanced learning algorithm, and good results were obtained. To class-imbalanced datasets evaluation, some reasonable model evaluation measures [17] including Precision, Recall and AUC/PRC (the area under Precision-Recall Characteristic curve) as model evaluation measure of VPD were utilized. Section 2 contains description of related data preprocessing algorithms used and includes the proposed models. Section 3 contains database for experimental setup and use. Section 4 illustrate the experimental results and analysis followed by Conclusion.

II. THE PROPOSED APPROACH

In this paper, we proposed methods consists of three parts. In the first stage, MFCC coefficients and their first-order and second-order differences were extracted from sustained vowels input signals, and the feature dimension is 36. In the second stage, imbalanced learning algorithm SMOTE, Borderline-SMOTE and ADASYN were applied to oversample the class-imbalanced initial dataset to obtain a class-balanced samples. In the third stage, voice pathology detection(VPD) models were contracted using support vector machine (SVM), decision tree (DT) and random forest (RF).

Three oversampling methods are combined with three common classifiers to prove the effectiveness of our approach. The diagram of the method is shown in Fig. 1.
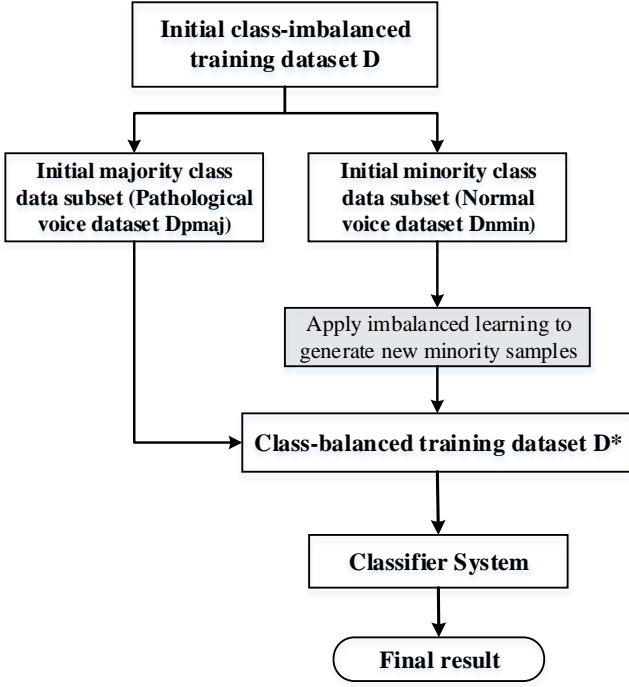
Fig. 1. The flow chart of VPD system based on imbalanced learning.

A. Feature Extraction

The well-known MFCCs feature extraction is widespread used in the voice pathology detection [18]. The Mel scale frequency domain extracted by MFCCs contains the non-linear characteristics of human ear frequencies. Fig. 2 shows the generalized block diagrams of MFCCs feature parameters extraction processes.
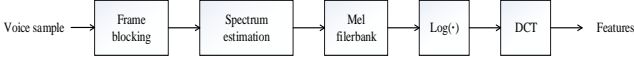
Fig. 2. Block diagram of MFCCs features extraction.

B. Data Preprocessing

- SMOTE as one of the common oversampling methods to deal with class-imbalanced datasets, the synthetic minority over-sampling technique (SMOTE) artificially generates new minority class samples and combines them with the initial training set to balance the training dataset. SMOTE generates synthetic samples of three steps. Firstly, it randomly selects a minority $\vec{a}$ to observe. Then, sample $\vec{b}$ is selected from the k nearest minority samples neighbors of sample $\vec{a}$. Finally, they randomly linearly interpolate the two samples to create a new sample $\vec{x}$, and the $\vec{x}$ function is shown in (1)

$$\vec{x} = \vec{a} + w \times (\vec{b} - \vec{a}) \qquad (1)$$

where w is the random weight in [0, 1]. A sample example of SMOTE is illustrated in Fig. 3b, where circular points are the minority class, square points are the majority class.

- Borderline-SMOTE identifies the sample belonging to boundaries via comparing the majority to minority class within each class neighborhood that needs to be oversampled, and generates new minority samples by incorporating borderline information, which expands decision space for the minority. A sample example of Borderline-SMOTE is shown in Fig. 3c, where circular points are the minority class, square points are the majority class.

- Adaptive synthetic sampling (ADASYN) algorithm main idea proceeds from the weighted distribution assumption based on the level of difficulty of each minority samples, according to which different numbers of synthetic samples are created adaptively. ADASYN is provided in Fig. 3d, where circular points are the minority class, square points are the majority class.

(a)   (b)

(c)   (d)
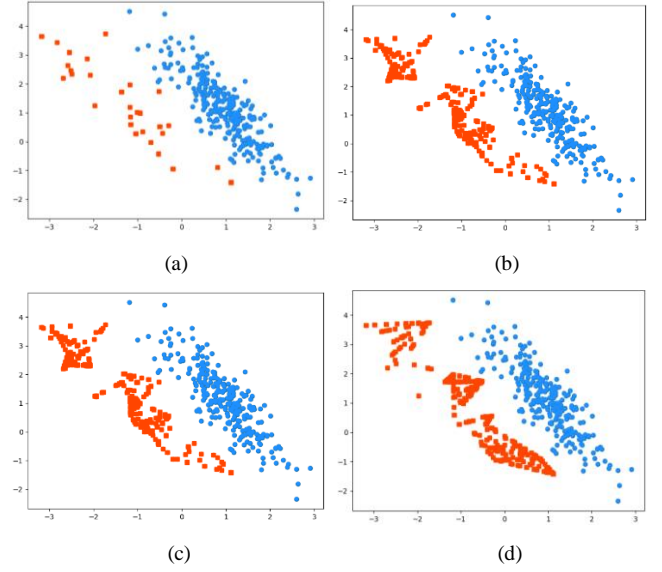
Fig. 3. Imbalanced learning synthesizes new samples. (a) is the initial dataset distribution. (b) is the dataset distribution after SMOTE treatment. (c) is the dataset distribution after Borderline-SMOTE treatment. (d) is the dataset distribution after ADASYN treatment.

C. Proposed Models

- Support Vector Machine (SVM) classifies data through the statistical concept of support vector, and it could be well extended to multiclass. SVM constructs linear or non-linear hyperplane for decision making to perform classification. An optimal hyperplane is defined as farthest from the nearest training data point of any class.

- Decision Tree (DT) is a tree-like classifier, with internal nodes and branches representing features and decision rules, respectively. DT is an example-based concluding study method, which does not depend on the assumption of probability distribution and is divided in a recursive manner.

- Random Forest (RF) is a combined classification algorithm proposed by Breiman. RF is classified by creating multiple classification and regression trees, each trained in the initial training data sample and finds on a randomly selected subset of input variables to determine segmentation.

III. Experimental Setup

A. Database

The Massachusetts Eye and Ear Infirmary (MEEI) collected the database we used. Since the establishment of the database, the MEEI database has been used in many research works. It is the most extensive and available voice disorder database [19]. Although the MEEI database is the most highly applied international commercial database in voice pathology detection, it has some limitations. Class-imbalanced distribution of samples is one of the major issues of the database. The MEEI database contains more than 1400 persistent vowel /a / speech files from 657 pathological speakers, but only 53 normal voice samples by normal speakers. The duration of all signals is 1 to 3 seconds, and the sampling frequency is 50 kHz or 25 kHz. In this study, all available 53 normal samples and four pathological types were selected, namely vocal cord nodules, vocal cord polyps, vocal cord edema, and vocal cord paralysis, as these four are common in clinical diagnosis. The number and gender distribution of samples are shown in Table I.

TABLE I.   NORMAL AND PATHOLOGICAL VOICE SAMPLES FORM MEEI (NO. OF MALE SPEAKERS + NO. OF FEMALE SPEAKERS)

Normal	Pathological					Total
	Nodules	Polyps	Paralysis	Edema	Total	
53 (21+32)	19 (1+18)	20 (12+8)	67 (37+30)	43 (32+11)	149	202

The number of voice samples from the MEEI database in the tables above can reflect that class-imbalanced distribution between normal and pathological voice samples on MEEI database. However, Class-imbalanced distribution of samples will lead to poor performance of classifiers, which has been ignored by many researchers in the past.

B. Parameter Settings

All 50 kHz files were down-sampled to 25 kHz to ensure that all voice files have the same sampling frequency. MFCCs parameters together with their first and second derivatives were extracted from each the file on MEEI database with a feature dimension of 36. The optimal parameter values of classifier were selected by using the grid search, the RBF kernel was selected for SVM, the penalty coefficient and the kernel parameter were 2 and 0.027. Entropy was chosen as a criterion in DT. The number of trees in the RF was 500, and the maximum depth of the RF was 13. Imbalanced learning algorithms all used default parameters set in Python's imbalanced-learn package. Throughout the experiment, 10-fold cross validation method of testing is used.

C. Model Evaluation Measure

Traditionally, the commonly used metrics to measure classifier performance are accuracy (Acc) and error rate (Err). However, the usual measure of classifier performance in the imbalanced datasets is inappropriate. If the wrong metric is chosen to evaluate models, it is likely to choose a poor model. To solve this issue, several reasonable metrics including Precision, Recall and PRC curves as model evaluation measure. Precision, Recall and the AUC (area of the PR curve) are calculated from (2) to (4). Table II shows the meanings of TP, TN, FP and FN. In the confusion matrix, TP is the number of true positive samples (actual minority, and predicted as minority), FP is the number of false positive samples (actual majority, but predicted as minority), FN is the number of false negative samples (actual minority, but

predicted as majority), TN is the number of true positive samples (actual majority, and predicted as majority).

TABLE II.   NORMAL AND PATHOLOGICAL VOICE SAMPLES FORM MEEI (NO. OF MALE SPEAKERS + NO. OF FEMALE SPEAKERS)

Actual classes	Prediction results	
	Positive class (Normal)	Negative class (Pathological)
Positive class (Normal)	TP	FN
Negative class (Pathological)	FP	TN

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

$$AUC = \sum_n (Recall_n - Recall_{n-1}) \times Precision_n \qquad (4)$$

D. Results and Analysis

Initial class-imbalanced dataset (53 normal voices and 149 pathological voices) is processed with imbalanced learning techniques by three oversampling algorithms. The generated class-balanced dataset contain 149 normal and 149 pathological voice samples, respectively. Classifier models were designed for each oversampling algorithm for evaluation. The model evaluation measures for the three classifier models designed with the initial class-imbalanced dataset and the three classifier models is designed using imbalanced learning are recorded in Table III.

TABLE III.   VOICE PATHOLOGY DETECTION RESULTS OF MODEL EVALUATION MEASURE

	SVM		DT		RF	
	Precision	Recall	Precision	Recall	Precision	Recall
Initial dataset	0.97	0.85	0.91	0.84	0.98	0.90
SMOTE	0.99	0.99	0.95	0.94	1.00	1.00
Borderline-SMOTE	0.97	0.97	0.97	0.97	1.00	1.00
ADASYN	0.97	0.97	0.93	0.93	1.00	1.00

In our experiments, Precision and Recall are measurements of exactness and predicting the completeness of minority class samples, respectively. Different classifier models show little difference performance on Precision and Recall values without imbalanced learning. According to Table III, the three classifier models combined with imbalanced learning have significantly improved Precision and Recall values. In detail, SVM obtains the highest mean Precision and Recall values with SMOTE, model increase the mean Precision from 0.97 to 0.99, increase the mean Recall from 0.85 to 0.99. SVM has the same Precision and Recall values when combined with Borderline-SMOTE and ADASYN, both of which are 0.97. For the DT, when the DT is combined with Borderline-SMOTE, the model shows the best results in Precision and Recall value, which are 0.97 and 0.97, respectively. Precision and Recall values increase for SMOTE-DT from 0.91 and 0.84 respectively to 0.95 and 0.94, and for ADASYN-DT from 0.91 and 0.84 respectively to 0.93 and 0.93. For the ensemble model RF, although it shows the same performance in Precision and Recall values with other models in the class-imbalanced dataset, it can be well combined with each oversampling algorithms, and the

Precision and Recall values trained in the balanced data set are improved from 0.98, 0.90 to 1.00, and 1.00, respectively. In summary, the above results mean that in pathological voice detection model, using imbalanced learning algorithm can greatly raise the detection accuracy of the model for a minority class samples.

We trained and tested the AUC/PRC of each classifier model before and after three imbalanced learning algorithms treatment, as shown in Fig. 4, and evaluated the model using AUC values.
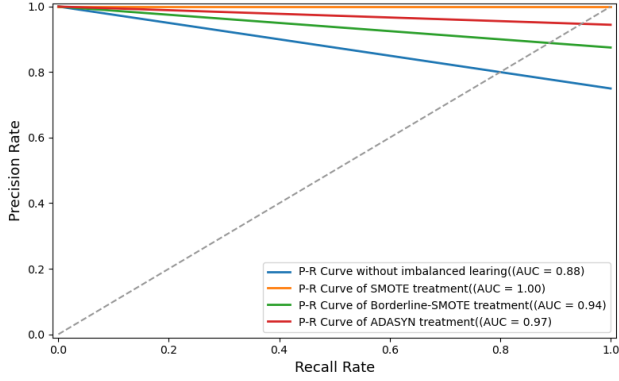
(a) PRC curve for SVM
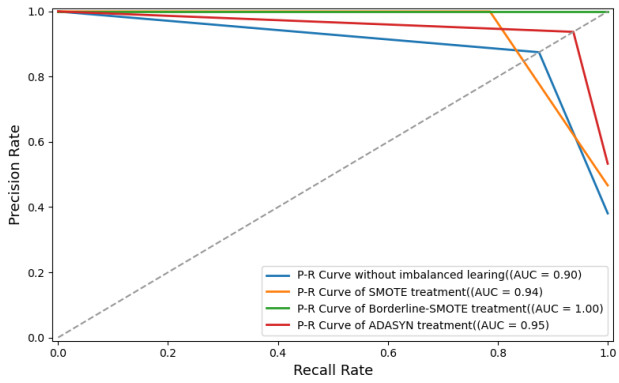
(b) PRC curve for DT
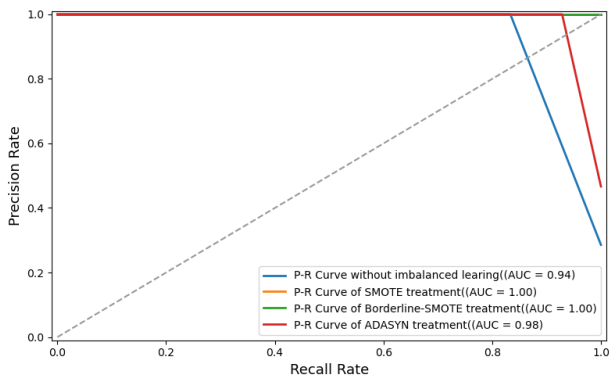
(c) PRC curve for RF

Fig. 4. PRC curves for different classifier models.

In case of class-imbalanced data, PRC curve can provide more information and reflect the performance of classification model and is less affected by the imbalanced data distribution. Precision and Recall rate are contradictory variables. When the performance of the classifier changes greatly on FP, the $FP_{rate}(\frac{FP}{(FP+TN)})$ will not change much. Therefore, the PRC

curve has good performance when the data is imbalanced. As for the AUC results, the VPD system combined with imbalanced learning algorithm shows high AUC value of each classifier models. SVM and DT are combined with SMOTE and Borderline-SMOTE, respectively, the value of the classifier model is the highest, with a value of 1.00. For RF, although ADAYSN-RF does not get 1.00 on the AUC metric, it also shows good performance, because RF as an ensemble classifier model is better than a single classifier in handling class-imbalanced data.

E. Comparison Existing Works

TABLE IV. COMPARISON WITH OTHER WORKS

Methods	Feature Name	Classifier	%SE	%SP
[18]	MFCCs	SVM	100.00%	46.87%
[20]	MDVP	FDR	93.22%	45.20%

Literatures [20] and [18] in Table IV, the subsets of "53 normal/93 pathological voices" and "53 normal/173 pathological voices" on the MEEI database was used, respectively. The accuracy of pathological voice detection model in literatures [20] and [18] was 76.36% and 94.07%. However the Sensitivity, which can measures the classifier's capability to recognize pathological samples (majority class) were 93.22% and 100.00%, and the Specificity used to represent the classifier's ability to recognize normal samples (minority class) were 45.20% and 46.87%. Meanwhile, it also indicates that training a model in an imbalanced dataset is misleading to choose Acc as the model evaluation measure. The pathological voice detection model proposed in this paper uses imbalanced learning, although it sacrifices a little recognition ability for a majority class samples, each classifier model greatly improves the recognition ability for a minority class samples compared with the literature [20] and [18].

IV. CONCLUSION

In order to improve the classification and generalization performance of VPD models on class-imbalanced databases, we propose a modeling method for pathological voice detection using imbalanced learning algorithm on MEEI database and reasonable model evaluation measures. This paper demonstrates the significance of the imbalanced algorithm in the case of a VPD model is established based on MEEI database. The presented method was tested with three classifier models SVM, DT and RF using 10 fold cross validation. The best classifier model is RF, which shows perfect performance on both Precision, Recall values and PRC curves after combining three imbalanced learning algorithms. SVM and DT models combined with SMOTE and Borderline-SMOTE showed the highest Precision, Recall and AUC values, respectively. The proposed models were compared with previous models showing greatly improved the recognition ability of a minority class samples. The results show that using imbalanced learning, the classifier models performed good in recognizing the minority class and using ensemble classifiers may be further improved pathological voice detection model.

REFERENCES

[1] Vikram, C.M. and Umarani, K., 2013, August. Phoneme independent pathological voice detection using wavelet based MFCCs, GMM-SVM hybrid classifier. In 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 929-934). IEEE.

[2] M.M. Constantine, C. Kotropoulos, I. Pitas, N. Maglaveras, Automatic detection of vocal fold paralysis and edema, IEEE Trans. Sys. Man Cybern. Part B 6 (November 2004) 31–35.Y. Zhang, J. Cao, W.-H. Liao, L. Zhao, J. Lin, "Theoretical modeling and experimental verification of circular Halbach electromagnetic energy harvesters for performance enhancement," Smart Materials and Structures, vol. 27, no. 9, pp. 095019, Aug. 2018.

[3] Kay Elemetrics Corp. Disordered Voice Database (CD-ROM), Version 1.03. Boston, MA: Massachusetts Eye and Ear Infirmary (MEEI), Voice and Speech Lab; 1994.

[4] Q. Wang, Y. Tian, and D. Liu, "Adaptive fh-svm for imbalanced classification," IEEE Access, vol. 7, pp. 130 410–130 422, 2019.

[5] Z. Sun, Q. Song, X. Zhu, H. Sun, B. Xu, and Y. Zhou, "A novel ensemble method for classifying imbalanced data," Pattern Recognition, vol. 48, no. 5, pp. 1623–1637, 2015.

[6] Cheng F , Zhang J , Wen C . Cost-Sensitive Large margin Distribution Machine for classification of imbalanced data[J]. Pattern Recognition Letters, 2016, 80(sep.1):107-112.

[7] Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H. and Bing, G., 2017. Learning from class-imbalanced data: Review of methods and applications. Expert Systems with Applications, 73, pp.220-239.

[8] Tahir, M.A., Kittler, J., Mikolajczyk, K. and Yan, F., 2009, June. A multiple expert approach to the class imbalance problem using inverse random under sampling. In International workshop on multiple classifier systems (pp. 82-91). Springer, Berlin, Heidelberg.

[9] Yeh, C.W., Li, D.C., Lin, L.S. and Tsai, T.I., 2016, July. A learning approach with under-and over-sampling for imbalanced data sets. In 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI) (pp. 725-729). IEEE.

[10] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," Journal of artificial intel-ligence research, vol. 16, pp. 321–357, 2002.

[11] H. Han, W.-Y. Wang, and B.-H. Mao, "Borderline-smote: a new over-sampling method in imbalanced data sets learning," in International conference on intelligent computing. Springer, 2005, pp. 878–887.

[12] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in 2008 IEEE international joint conference on neural networks (IEEE world congress on compu-tational intelligence). IEEE, 2008, pp. 1322–1328.

[13] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik, "Support vector clustering," J. Mach. Learn. Res., vol. 2, pp. 125–137, Mar. 2002.

[14] Quinlan J R . Induction of decision trees[J]. Machine Learning, 1986, 1(1):81-106.

[15] Breiman L. Random forests. Machine learning. 2001 Oct 1;45(1):5-32.

[16] Marina Sokolova, Nathalie Japkowicz, Stan Szpakowicz. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation[J]. Lecture Notes in Computer ence, 2006, 4304:1015-1021.

[17] J.I. Godino-Llorente, P. Gomes-Vilda, M. Blanco-Velasco, Dimension-ality reduction of a pathological voice quality assessment system based on Gaussian mixture models and short-term cepstral parameters, IEEE Trans. Biomed. Eng. 53 (October (10)) (2006) 1943–1953.

[18] N. Saenz-Lechon, J.I. Godino-Llorente, V. Osma-Ruiz, P. Gomez-Vilda, Methodological issues in the development of automatic systems for voice pathology detection, Biomed. Signal Process. Control 1 (2) (2006) 120–128.

[19] A. Al-Nasheri, G. Muhammad, M. Alsulaiman, Z. Ali, T. A. Mesallam, M. Farahat, K. H. Malki, and M. A. Bencherif, "An investigation of multidimensional voice program parameters in three different databases for voice pathology detection and classification," Journal of Voice, vol. 31, no. 1, pp. 113–e9, 2017.

Multi-label Feature Selection for Long-term Electrocardiogram Signals

Yuwen Li
School of Instrument Science and Engineering
Southeast University
Nanjing, China
liyuwen@seu.edu.cn

Zhimin Zhang*
Science and Technology on Information Systems Engineering Laboratory
The 28th Research Institute of CETC
Nanjing, China
zmzsdu@163.com

Fan Zhou
School of Instrument Science and Engineering
Southeast University
Nanjing, China
fanz1996@seu.edu.cn

Yantao Xing
School of Instrument Science and Engineering
Southeast University
Nanjing, China
230198304@seu.edu.cn

Jianqing Li
School of Biomedical Engineering and Informatics
Nanjing Medical University
Nanjing, China
ljq@seu.edu.cn

Chengyu Liu*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
chengyu@seu.edu.cn

Abstract—**Arrhythmia is a kind of cardiovascular disease that seriously threats human health, and intelligent analysis of electrocardiogram (ECG) is an effective method for the early prevention and precise treatment to arrhythmia. In clinical ECG databases, it is common to see the multi-label phenomenon that one patient would be labelled with multiple types of arrhythmia. However, the current research is mainly to use the multi-class methods for dealing with the problem of multi-label, ignoring the correlations between diseases and causing information loss. Therefore, this paper aims to propose a multi-label feature selection (MLFS) method based on ECG and design a novel evaluation criterion based on kernelized fuzzy rough sets so as to choose the optimal feature subset and optimize ECG feature space. Through sufficient experiments to prove the feasibility of our methods, we obtain the optimal feature subset composed of 23 ECG features. For the six evaluation criterions of multi-label learning, Average Precision is 0.8053, Hamming Loss is 0.1063, Ranking Loss is 0.1366, One-error is 0.2021, Coverage is 0.4018, and Micro-F1 is 0.5874. The outcome presents great superiority to the current algorithms of MLFS. This study is a prerequisite for implementing big data ECG classification diagnosis and disease modeling.**

Keywords—Biomedical information processing, ECG signals, Arrhythmia, Multi-label learning, Feature selection

I. INTRODUCTION

According to the World Health Statistics 2018 [1] reported by World Health Organization (WHO), cardiovascular disease is still the main cause of death in the world. Cardiovascular disease has become "the first killer" to threaten human life and health. Arrhythmia is one of the most common and extremely high incidence cardiovascular diseases, which mainly causes palpitations, chest tightness, dizziness, hypotension and other symptoms. In severe cases, syncope or even sudden death may occur. The majority of arrhythmias can be clearly diagnosed through long-term continuous electrocardiogram (ECG) monitoring. The early intervention of the causes that may lead to serious consequences can significantly reduce the death/disability rate of cardiovascular diseases and greatly reduce social and economic losses. On the other hand, long-term ECG signals obtained through monitoring can be used to catch abnormal cardiac beats in time to avoid missing detection of accidental or paroxysmal arrhythmias [2]. Therefore, through continuous monitoring and automatic

analysis of long-term ECG signals, timely and rapid detection of arrhythmia, and classification of arrhythmia types, can not only help doctors improve work efficiency, but also is the key to early prevention of arrhythmia.

Under the support of the technology of "Internet + intelligent medical treatment" in the new era, Artificial Intelligence (AI) technology has promoted the development of a new mode of arrhythmia prevention and control. In recent years, AI technology has made great breakthroughs in the automatic diagnosis of ECG signals [3-5] to help improve early detection accuracy, strengthen diagnosis and risk control, reduce treatment costs, assist patients in self-health management, and improve treatment effect.

Machine Learning (ML) as a way to realize AI, is used for intelligent classification and diagnosis of arrhythmias. It refers to the process of automatic classification and disease diagnosis of ECG signals after the detection and quantification of ECG signal feature. Researchers at home and abroad hope to use ML algorithm to assist doctors to complete automatic diagnosis and analysis of ECG signals with high accuracy and high speed. Therefore, it is an urgent problem to build an intelligent analysis model for arrhythmias from the perspective of improving the accuracy and robustness of ECG classification by means of ML.

Although the development of ML algorithm has brought new research content to the traditional topic of intelligent diagnosis algorithm for arrhythmia, there are two important problems to be solved: (1) How to solve the problem of multi-label with ECG data? In actual clinical cases, arrhythmias are caused by abnormalities in the frequency, rhythm, pacing site, conduction speed and excitation sequence of cardiac impulses. The changes of cardiac electrical signals in cardiovascular diseases are varied, leading to overlapping and mutual transformation of various arrhythmias. Thus, in clinical ECG data, one patient may correspond to multiple arrhythmia types. This is known as "multi-label phenomenon". Multi-label problem is common in clinical ECG database. (2) How to deal with the high-dimensional problem of ECG features? When ECG signals are mapped to the feature space, they often still have high dimension. Feature selection in high-dimensional data space is always a challenge in ECG processing. In addition, the original ECG features obtained by different ECG feature extraction methods may be redundant or irrelevant for the arrhythmia classification task.

II. RELATED WORK

A. Feature Selection of ECG Signals

Feature selection is a common and important dimensionality reduction method in ML. Its purpose to select the optimal feature subset with some features from feature space through certain criteria. Appropriate feature selection method can reduce the number of features, eliminate irrelevant, redundant or noise data, improve training speed and performance, and improve prediction accuracy [6].

With ensuring the accuracy of classification, the filter feature selection for ECG signals is aiming at selecting the relevant and indispensable features from the original ECG feature set to form the optimal feature subset. It should not only be able to express the original model to a certain extent, but also minimize the loss of information. Kamath et al. [7] presented a method based on energy operators for feature extraction, and the classification accuracy of 67,960 heart beats was 95%. Shen et al. [8] put forward an adaptive feature selection algorithm using wavelet coefficient, which could improve the accuracy of heart beat classification from 80.32% to 98.92%. However, the dimension of feature space was still greater than 50 after dimension reduction. Martis et al. [9] compared the effect of three feature selection methods including principal component analysis, independent component analysis and linear discrimination, and verified the five-classification effect of heart beat by probabilistic neural network.
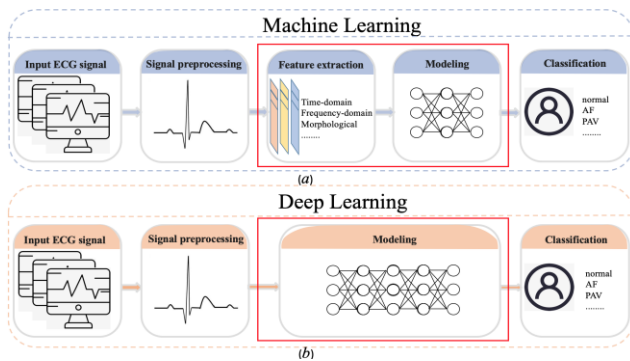
Fig. 1. Traditional machine learning and deep learning modeling steps for intelligent classification of ECG signals.

To sum up, the existing researches mainly focus on the feature selection of single-label ECG signals. However, the feature selection algorithm based on ECG signals with "multi-label phenomenon" has not been clearly reported. Therefore, we propose the feature selection algorithm for multi-label ECG data to explore the redundancy between features. At the same time, we also mine the intrinsic correlation between features and labels or label set under the many-to-many mapping relationship. Multi-label feature selection (MLFS) aims at selecting some discriminative features among all multi-label ECG features. Therefore, we construct feature selection for multi-label ECG data can reduce computational complexity and build up the learning performance.

B. Analysis of Arrhythmias Based on ML

The results of arrhythmia analysis and classification algorithm based on ML are various. However, these algorithms are quite different in the aspects of learning and recognition process of ECG features, and the performance and application direction of these algorithms are also different. The existing algorithms can be roughly divided into two

categories: arrhythmia classification based on traditional ML and Deep Learning (DL).

The classification method of arrhythmias based on traditional ML uses various ECG features to acquire the ability of disease classification. It has strong flexibility and good effect of data classification. The modeling steps are shown in Fig. 1a. Polat et al. [10] used the least squares support vector machine (LSSVM) method to classify abnormal ECG signals. The algorithm can distinguish the normal heart rate from the abnormal heart rate. Melgani et al. [11] presented a classification method of ECG based on hidden markov tree, which can distinguish 8 types of arrhythmias and continuously adjust the states and the percentage of gauss in the model in the process of recognition. Krimi et al. [12] improved the generalization performance of the classification of heartbeats using support vector machine (SVM) through particle swarm optimization classification algorithm.

The modeling steps of the classification method of arrhythmias based on DL are shown in Fig. 1b. In 2017, Acharya et al. [13] researched the arrhythmias detection using an 11-layer one-dimensional convolutional neural network structure. In 2018, Faust et al. [14] used two-way LSTM (Long Short Term Memory) neural network to identify AF. In 2018, Wang et al. [15] put forward the recursive neural network method and used it to find supraventricular ectopic and ectopic ventricular beat. In 2018, Xia et al. [16] applied the convolutional neural network to the detection of AF for ECG signals. In 2019, Andrew Y. Ng et al. [17] published a study in Nature Medicine. Ng's team proposed a deep learing model about the residual neural network for accurately detecting 10 types of arrhythmia, sinus rhythm and noise by a single lead ECG. It can be seen from the literature survey that the classification model of arrhythmias based on DL has excellent performance. However, it is widely criticized for its lack of interpretability due to the parameter sharing and complex feature extraction and combination within the DL model. The task of arrhythmia analysis is not only the superiority of classification accuracy, but also to assist doctors to understand the ECG features and understand more transparently why the classification model makes such decisions and what features play an important role in the decision. Therefore, it is very important to study multi-label feature learning for the analysis of arrhythmias.

In the past few decades, most of the common MLFS algorithms are to convert multi-label data into multiple single-label data, so that each task corresponds to one label. This strategy is a simple and straightforward solution. However, it has two obvious disadvantages when dealing with the multi-label problem: (1) information loss and resource waste. Some multi-label information is bound to be ignored. However, all information of labels provided by ECG experts is helpful for ECG analysis, and the labels of ECG signals is expensive, so it is bound to cause information loss and resource waste. (2) this strategy ignores the correlation between all types of arrhythmias since all types of arrhythmias are correlated.

In this paper, for ECG signal, a MLFS algorithm based on kernelized fuzzy rough sets is design to construct the standard and universal optimal feature optimization method and select the features with the most information. Moreover, kernelized fuzzy dependency is defined to determine the importance of features. ECG feature space is optimized through the selection of multi-label features to provide effective features for the accurate identification of arrhythmia, so as to provide

theoretical and technical support for the early prevention, long-term monitoring and diagnostic evaluation of cardiovascular diseases.

III. DATABASE

An independent, open access dataset from 11 different hospitals was provided by the China Physiological Signal Challenge (CPSC) [18]. This database contained two parts: CPSC 2018 Training Set (open online) and CPSC 2018 Test Set (unopen). 12 leads ECG recordings sampled as 500 Hz last from 6s to just 60s. This data set is multi-label data that includes nine types of ECG signals. However, a few recordings have up to three labels.

This article adopts CPSC 2018 Training Set as training data set. CPSC 2018 Training Set contains 6,877 recordings. This article only selects the patients with normal type or 5 types of common arrhythmia (atrial fibrillation, right bundle branch block, left bundle branch block, atrial premature beat, ventricular premature beat). These 5 abnormalities all have the characteristics of high prevalence. Finally, we choose 5078 recordings with the first lead.

This paper also adopts CPSC 2018 Test Set as test data set. CPSC 2018 Test Set contains 2,954 recordings. This article chooses 2175 recordings with 6 types of ECG signals.

IV. METHODS

A. Feature Extraction

Before feature extraction, the QRS-wave, P-wave, and T-wave positions of all ECG signal are first detected. In this paper, we use the R-wave position detection algorithm [19-20]. Then we apply the concept of refractory period and backtracking technology to effectively avoid R-wave missing detection, while ensuring that ECG signals with very low amplitude and no change are not detected [21-22]. In addition, ECG signal with less than 12 RR intervals were excluded. Then, the algorithm proposed by Shang et al. [23-24] is used to detect the position of T-wave, Q-wave, S-wave, P-wave and R-wave, and the open source algorithm prepared by Datta et al. [25].

In this paper, 117 features are extracted according to the position of P-wave, Q-wave, R-wave, S-wave and T-wave. These features are divided into four types, 27 time-domain features, 34 frequency-domain features, 30 morphological features and 26 nonlinear features.

B. Problem Description

According to the above feature extraction methods for ECG signals, 117 ECG features $\mathcal{C} = \{f_1, f_2, ..., f_{117}\}$ were obtained. We suppose that the ECG signal sample set $\mathcal{A} = \{a_1, a_2, ..., a_n\}$ has 6 arrhythmia labels constructed the arrhythmia label set $\mathcal{Y} = \{y_1, y_2, ..., y_6\}$. In the training set $\mathcal{D} = \{(x_i, y_i)|1 \leq i \leq n\}$, $x_i \in \mathbb{R}^{117}$ is a 117-dimension feature vector, and $y_i \in \{0,1\}^6$ is 6-dimension binary label vector. If a_i has the label y_j, then $y_{ij} = 1$ for the vector y_i, otherwise $y_{ij} = 0$.

C. MLFS for ECG

Firstly, for feature space, regarding each ECG feature f_i, we use Gaussian kernel function [26] for evaluate the ECG feature have that makes them like each other ECG signals, and the combined strategy is to minimum. For label space, considering each label y_i, we use Match kernel function [26]

for analyze the arrhythmias categories of ECG signals sharing labels of knowledge, and the combined strategy is to take the sum. Combination of this two spaces is shown in Fig. 2.
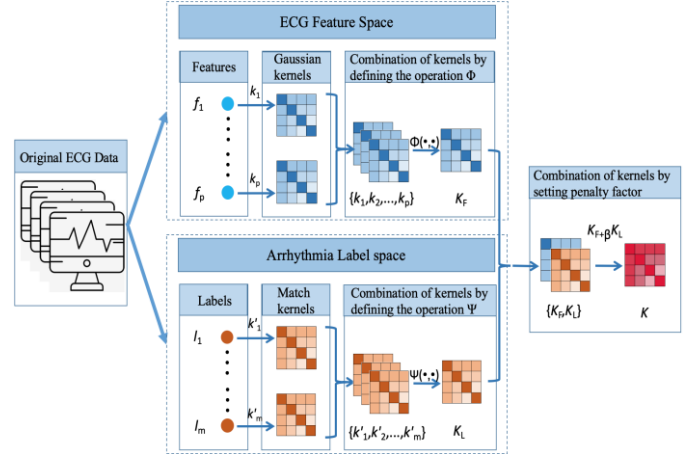
Fig. 2. Workflow of multiple kernel learning for ECG data.

1) Combination strategy of two space:

Definition 1. Let $MD = <\mathcal{A}, \mathcal{C} \cup \mathcal{Y}>$ be a multi-label decision system. K_S is kernel matrix about feature subset $S \subseteq \mathcal{C}$, and K_L is kernel matrix in label space L. Let $\delta \in [0,1)$ be the penalty factor. The combination of kernel matrix $K^{\{S,L\}}$ between feature subset B and label space L is defined as

$$K^{\{S,L\}} = K_S - \delta \widetilde{K}_L \quad (1)$$

In (1), $\widetilde{K}_L = K_L/6$. It is clear that all elements of K_S are in [0,1], while all elements of K_L are in [0,6]. Aiming to achieve dimensionless quantity, we use $\widetilde{K}_L = K_L/6$ to amend K_L, and $\widetilde{K}_L$ represents the label overlap ratio between samples in the label space L. $K^{\{S,L\}}$ indicates the probability of being the nearest different classes' sample. $O(x_i)$ and $P(x_i)$ are proposed to be sets of kernel different classes' samples and kernel same class's samples, respectively. $O(x_i)$ as well as $P(x_i)$ can be calculated by [27].

2) Construction Multi-label Kernelized Fuzzy Rough Set Model:

Definition 2. For $\mathcal{A}(x \in \mathcal{A})$, $O(x)$ and $P(x)$ are known. The lower and upper approximations are analyzed by

$$\underline{K^R}D(x) = \operatorname*{mean}_{u \in O(x)}\{1 - K_S(x,u)\} \quad (2)$$

$$\overline{K^R}D(x) = \operatorname*{mean}_{u \in P(x)}\{K_S(x,u)\} \quad (3)$$

3) The importance of ECG feature:

For $S \subseteq \mathcal{C}$, $\underline{K_S^R}(L)(x)$ is assumed to be the lower approximation of $x \in \mathcal{A}$. Then the multi-label fuzzy positive region $POS_S(L)$ can be evaluated as

$$POS_S^R(L) = \sum_{x \in \mathcal{A}} \underline{K_S^R}(L)(x) \quad (4)$$

By the above definition of fuzzy positive region, we can express the multi-label fuzzy dependency function as

$$\gamma_S^R(L) = \frac{POS_S^R(L)(x)}{|\mathcal{A}|} = \frac{\sum_{x \in \mathcal{A}} \underline{K_S^R}(L)(x)}{|\mathcal{A}|} \quad (5)$$

From the following equation (6), we can calculate the importance of the any feature $f \in \mathcal{C} - S$,

$$Sig^R(f, S, L) = \gamma_S^R(L) - \gamma_{S \cup f}^R(L) \qquad (6)$$

4) MLFS Algorithm:

The significance of each feature is measured by (6). In general, a selection algorithm contains two crucial factors, which are feature evaluation and search strategy. The forward greedy algorithm is applied as the search strategy. Additionally, Algorithm 1 presents the pseudo code of RMFRS.

Algorithm 1: RMFRS

Input: $\mathcal{C}$: a set of 117 ECG features; $\mathcal{A}$: a set of ECG recordings; L: the whole arrhythmia label space; β : a parameter.

Output: *red*: one reduct

1. $red \leftarrow \emptyset$

2. **for** each $f_i \in \mathcal{C} - red$ **do**

3.   **for** $x_i \in U$ **do**

4.   compute $K_{red \cup f_j}$, K_L and $K^{\{red \cup f_j, L\}}$

5.   find $O(x_i)$

6.   compute $\gamma_{red \cup f_j}^R(L)$

7.   compute $Sig^R(f, red, L) = \gamma_{red \cup f_j}^R(L) - \gamma_{red}^R(L)$

8.   **end for**

9.   find feature f_i with maximum value $Sig^R(f, red, L)$

10.   **if** $Sig^R(f, red, L) > 0$ **then**

11. $red \leftarrow red \cup f_j$

12.   **else**

13.   return *red*

14.   **end if**

15. **end for**

16. return *red*

V. Experiments

In order to make an assessment of RMFRS, the other three algorithms, namely MLNB [28], MDDM [29] and PMU [30] are compared with the analysis results. Moreover, to stay in consistency with these three counterpart algorithms, the training and testing set are also utilized. In the meantime, MLKNN (K=10) [31] is used to assess the classification performances of the algorithms. The specific experimental steps are as follows. Firstly, CPSC 2018 Training Set as training set, and 117 features were taken as the input of four MLFS algorithms. The output result is the optimal feature subset for each algorithm. Second, CPSC 2018 Test Set as test set, and the four kinds of feature subsets are respectively taken as the input of MLKNN. The output results are the occurrence probability corresponding to each label as the basis for label prediction. Finally, various metrics are used to evaluate the classification performance of the MLFS algorithms.

A. Training Setup

For the four MLFS algorithms, RMFRS and MLNB is able to attain the feature subset directly. Nevertheless, MDDM and PMU can only get the feature selection output by obtaining the list of features. Hence, in order to gain comparable results,

the feature rank results determines the number of features that will be selected to be the ultimate feature subset based on the length of feature subset with RMFRS algorithm. According to the suggestion in [28], two threshold parameters are carefully selected respectively, namely, the ratio is given as 0.3, and the smooth is set to be 1. The parameter is set as 0.5 for MDDM as referred by [29]. The values of β and μ are set from 0.1 to 0.5 with a step of 0.01 for our proposed RMFRS, so as to prove that the performance varies with β and μ. The experiments imply that when $\delta = 0.03$ and μ=0.38, the classification performance achieved the best for all evaluation metrics.

B. Evaluation Metrics

We ended up using both label set prediction and label ranking to evaluate the performance of the feature selection algorithms[32]. Concretely, Hamming Loss and Micro-F1 are selected in label set prediction, and Average Precision, Ranking Loss, Coverage, and One-error are selected in label ranking. As is known to all, these six criteria illustrate the performance comparisons from different aspects. On all these criteria, feature selection algorithms generally cannot be superior to other algorithms.

VI. Results

For the comparability of performances among all MLFS algorithms, the feature subset are fed to MLKNN as input. Fig. 3 shows the variation trend of various evaluation metrics with the number of selected features through the six evaluation criteria. We take Average precision, Hamming Loss, and Coverage as examples for show. In Fig. 3, horizontal axis of each graph indicates the number of the chosen features, while vertical axis of each graph indicates classification evaluation metrics. Moreover, four different colored lines represent MDDM, PMU, MLNB and RMFRS, respectively. The proposed algorithm RMFRS is the red line. As shown in Fig. 3, RMFRS can achieve the optimal classification performance with the increasing number of selected features. Notice that RMFRS has a certain number of features that makes it delivers better performance, and it is in line with the actual situation. RMFRS selects and ranks the importance of 23 features as follows: (1) normalized fuzzy entropy [33], (2) the ratio of the number of S waves less than Q waves to the number of S waves, (3) sample entropy [33] when embedding dimension is 1, (4) (the number of P wave)/ (the number of R wave), (5) sample entropy when embedding dimension is 2, (6) maximum of sample entropy, (7) delay correlation between RR interval and one sample sequences after itself [34], (8) maximum of approximate entropy [35], (9) dimension of maximum of sample entropy, (10) normalized low-frequency power, (11) percentage of ΔRR interval greater than 20ms, (12) normalized high-frequency power, (13) number of infinite sample entropy, (14) dimension of minimum of sample entropy, (15) minimum of RR interval, (16) root mean square of ΔRR interval, (17) radius of the smallest circle in the Lorentz chart which accounts for 60% of the area [36], (18) relevant parameters of Lorentz diagram [37], (19) polynomial coefficient fitted from least square, (20) variation coefficient of ΔRR interval, (21) coefficient of variation of S-wave slope, (22) the ratio of the number of ST segment slopes less than 0 to the number of R waves, (23) maximum of RR interval. These 23 ECG features include 4 features in time-domain, 3 in frequency-domain, 12 in nonlinear and 4 in morphological domain.
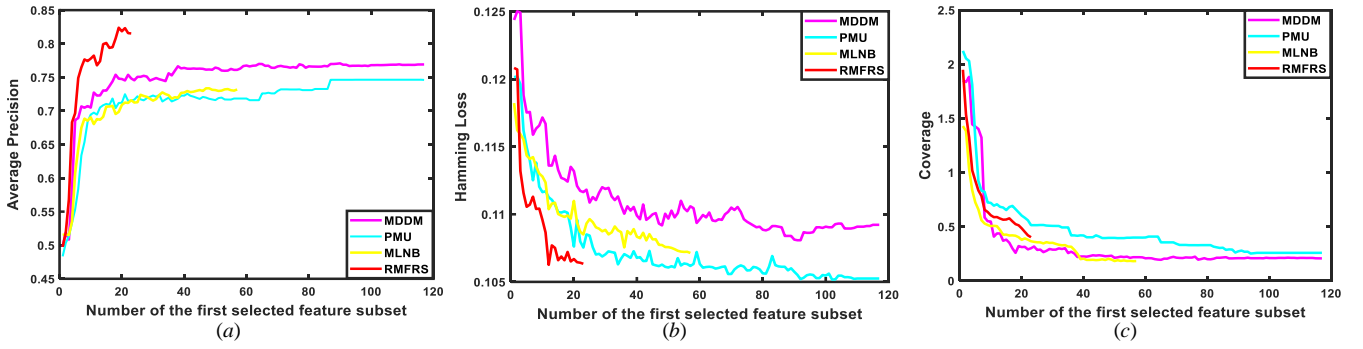
Fig. 3. Performance variations for the four MLFS algorithms based on (a) Average precision, (b) Hamming Loss, and (c) Coverage.

TABLE I. PERFORMANCE COMPARISON AMONG FOUR MLFS ALGORITHMS

Evaluation Metrics	MLFS Algorithms			
	MDDM	PMU	MLNB	RMFRS
Average Precision (↑)	0.7483	0.7155	0.7159	**0.8053**
Hamming Loss (↓)	0.1116	0.1075	0.1085	**0.1063**
Ranking Loss (↓)	0.1401	0.1523	0.1472	**0.1366**
One-error (↓)	0.2219	0.2342	0.2558	**0.2021**
Coverage (↓)	**0.3001**	0.5131	0.3646	0.4018
Micro-F1 (↑)	0.5621	0.5633	0.5125	**0.5874**

To prove the validity of RMFRS, a series of experiments are conducted which makes performance comparisons among RMFRS, MLNB, MDDM, PMU, and RF-ML. As is shown in Table I, various evaluation metrics are illustrated, with the optimal results being emphasized in bold. Moreover, "↓" means "the smaller the better", while "↑" conveys that "the larger the better", respectively. From Table I, the following conclusions can be easily drawn: (1) RMFRS outperforms other comparative algorithms in all evaluation indicators except Coverage. (2) Notice that the evaluation metrics of RMFRS is highly similar to that obtained by MDDM. These experimental results demonstrate that the proposed RMFRS is superior than the other existing feature selection algorithms.

VII. DISCUSSION

A. Analysis of Superiority

It is noted that even though the performance of the model is greatly improved after feature selection, RMFRS tends to have a better working effect than other comparing algorithms. The superiorities of RMFRS algorithm in feature selection of ECG signals come from two sources: (1) he fuzzy processing of ECG signals based on fuzzy rough set model. Even after the denoising filter, the extracted ECG features will be fuzzy and inaccurate to affect the ECG feature measurement and feature selection results. Therefore, the fuzzy rough set theory could be applied to do with the fuzziness and uncertainty of ECG features, so as to reduce their influence on feature evaluation. Then the fuzzy dependence evaluation function can be defined to scientifically and reasonably describe the correlation between ECG features and features as well as between features and arrhythmia labels. (2) The integrity of the arrhythmia label space is considered. Interestingly, the proposed method is superior to other existing MLFS methods in ECG dataset, which shows that the approximate plan given in this paper is more superior than that of other existing methods. For almost all conventional MLFS algorithms, only

the relationships within feature space and label space could be utilized. When it comes to RMFRS, our purpose is to gain additional information and build a stable approximation model from ECG feature space to arrhythmia label space. It is shown that the core information of arrhythmia labels is of great significance for MLFS.

B. Stability Analysis

Fig. 4 is drawn to show the results of verifying the stability of different MLFS algorithms for the six criteria. Due to the large differences in predictive classification performance using different evaluation metrics, the results in Fig. 4 are normalized to 0.1~0.5 as the stability indexes of different evaluation metrics. When the spider web diagram shows rounder, it means that the corresponding classification algorithm has higher the stability. Specifically, the stability value of RMFRS is represented by the red line. We can tell from Fig. 4 that RMFRS is superior to the other methods and performs stably on six evaluation criteria.
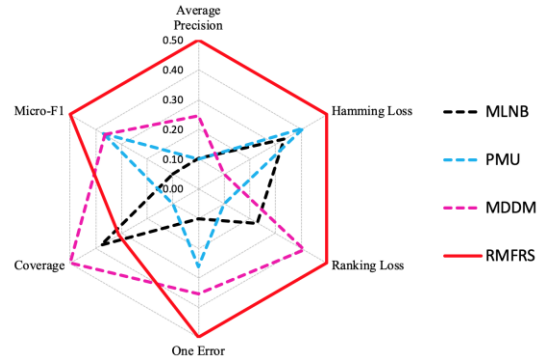
Fig. 4. Spider Web diagram of 6 different evaluation metrics.

VIII. CONCLUSION

ECG feature selection algorithm involves two aspects. On the one hand, it deals with high dimension. We can remove irrelevant features and redundant features in ECG features to avoid the loss of key information and "the curse of dimensionality". Besides, it deals with multi-label problem by illustrating the correlation between ECG features and arrhythmia categories. Therefore, it is a key scientific problem to explore the evaluation mechanism of feature importance in the multi-label ECG features. In this paper, a novel MLFS algorithm is proposed, i.e. RMFRS. Moreover, we define the fuzzy dependency to calculate the correlation between features as well as between features and the label set, and it was used as the evaluation criterion for the importance of features. Therefore, a consistent and non-redundant low-dimensional feature subset was constructed.

In this paper, only four commonly used MLFS algorithms are discussed. However, more algorithms including DL theory with increasing practical applications need to be verified. So the next step of work is to enrich database, choose different MLFS algorithm, and compare more model of classification algorithm to get deeper understanding for different feature set and classification of the role and influence. We will put forward a more targeted optimization to get more suitable for real-time, automatic and accurate detection of arrhythmia in dynamic ECG feature set and multi-label algorithm.

REFERENCES

[1] https://www.who.int/gho/publications/world_health_statistics/2018/en/.

[2] E. J. Topol. "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, no. 1, pp. 44-56, 2019.

[3] H. Blackburn, A. Keys, E. Simonson, P. Rautaharju, and S. Punsar, "The electrocardiogram in population studies-a classification system," *Circulation*, vol. 21, no. 6, pp. 1160-1175, 1960.

[4] C. Y. Liu, X. Y. Zhang, L. N. Zhao, et al, "Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 363-1374, 2019.

[5] G. H. Tison, J. Zhang, F. N. Delling, and R. C. Deo, "Automated and interpretable patient ECG profiles for disease detection, tracking, and discovery," *Circulation: Cardiovascular Quality and Outcomes*, vol. 12, no. 9, pp. e005289(1)-e005289(12), 2019.

[6] M. Dash, and H. Liu, "Feature selection for classification," *Intelligent Data Analysis*, vol. 1, no.3, pp.131-156, 1997.

[7] C. Kamath, "ECG beat classification using features extracted from teager energy functions in time and frequency domains," *IET Signal Processing*, vol. 5, no. 6, pp. 575-581, 2011.

[8] C. P. Shen, W. C. Kao, Y. Y. Yang, et al, "Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines," *Expert Systems with Applications*, vol. 39, no. 9, pp. 7845-7852, 2012.

[9] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and discrete wavelet transform," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 437-448, 2013.

[10] K. Polat, B. Akdemir, and S. Günes, "Computer aided diagnosis of ECG data on the least square support vector machine," *Digital Signal Processing: A Review Journal*, vol. 18, no. 1, pp. 25-32, 2008.

[11] F. Melgani, and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 5, pp. 667-677, 2008.

[12] S. Krimi, K. Ouni, and N. Ellouze, "ECG signal classification using hidden Markov tree," *International Review on Computers and Software*, vol. 5, no. 6, pp. 615-619, 2009.

[13] U. R. Acharya, H. Fujita, O. S. Lih, et al, "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network," *Information Sciences*, vol. 405, pp. 81-90, 2017.

[14] O. Faust, A. Shenfield, M. Kareem, et al, "Automated detection of atrial fibrillation using long short-term memory network with RR interval signals," *Computers in Biology and Medicine*, vol. 102, pp. 327-335, 2018.

[15] G. J. Wang, C. S. Zhang, Y. P. Liu, et al, "A global and updatable ECG beat classification system based on recurrent neural networks and active learning," *Information Sciences*, vol. 501, pp. 523-542, 2018.

[16] Y. Xia, N. Wulan, K. Q. Wang, and H. G. Zhang, "Detecting atrial fibrillation by deep convolutional neural networks," *Computer in Biology and Medicine*, vol. 93, pp. 84-92, 2018.

[17] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, et al, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, , vol. 25, no. 1, pp. 65-69, 2019.

[18] F. F. Liu, C. Y. Liu, L. N. Zhao, et al, "An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection," *Journal of Medical Imaging and Health Informatics*, vol. 8, no. 7, pp. 1368-1373, 2018.

[19] J. Behar, A. Johnson, G. D. Clifford, et al, "A comparison of single channel fetal ECG extraction methods," *Annals of Biomedical Engineering*, vol. 42, no. 6, pp. 1340-1353, 2014.

[20] J. Behar, J. Oster, and G. D. Clifford, "Combining and benchmarking methods of foetal ECG extraction without maternal or scalp electrode data," *Physiological Measurement*, vol. 35, no. 8, pp. 1569-1589, 2014.

[21] J. Pan, and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering,* vol. 32, no. 3, pp. 230-236, 1985.

[22] T. Sharma, and K. K. Sharma, "QRS complex detection in ECG signals using locally adaptive weighted total variation denoising," *Computers in Biology and Medicine*, vol. 87, pp. 187-199, 2017.

[23] H. X. Shang, S. S. Wei, F. F. Liu, et al, "An improved sliding window area method for T wave detection," *Computational and Mathematical Methods in Medicine*, pp. 1-11, 2019.

[24] A. A. Suárez-León, C. Varon, R. Willems, et al, "T-wave end detection using neural networks and support vector machines," *Computers in Biology and Medicine*, vol. 96, pp. 116-127, 2018.

[25] S. Datta, C. Puri, A. Mukherjee, et al, "Identifying normal, AF and other abnormal ECG rhythms using a cascaded binary classifier," *in 2017 Computing in Cardiology*, 2017.

[26] B. Schlkopf, and A. Smola, "Learning with kernels," *in: Proceeding of the 21th International Conference on Machine Learning*, 2001, pp. 639-646.

[27] Y. W. Li, Y. J. Lin, J. H. Liu, et al, "Feature selection for multi-label learning based on kernelized fuzzy rough sets," *Neurocomputing*, vol. 318, pp. 217-286, 2018.

[28] M. L. Zhang, J. M. P. Na, and V. Robles, "Feature selection for multi-label naive Bayes classification," *Information Sciences*, vol. 179, pp. 3218-3229, 2009.

[29] Y. Zhang, and Z. H. Zhou, "Multilabel dimensionality reduction via dependence maximization," *ACM Transaction Knowledge Discovery Data,* vol. 4, pp. 1-21, 2010.

[30] J. Lee, and D. Kim, "Feature selection for multi-label classification using multivariate mutual information," *Pattern Recognition Letter*, vol. 34, pp. 349-357, 2013.

[31] M. L. Zhang, Z. H. Zhou, "ML-KNN: a lazy learning approach to multi–label learning," *Pattern Recognition*, vol. 40, pp. 2038-2048, 2007.

[32] M. Semelka, J. Gera, and S. Usman, "Sick sinus syndrome: a review," *American Family Physician*, vol. 87, no. 10, pp. 691-696, 2013.

[33] C. Y. Liu, O. Julien, R. Erik, et al, "A comparison of entropy approaches for AF discrimination," *Physiological Measurement*, vol. 39, no. 7, pp. 74002(1)-74002(18).

[34] M. Bsoul, H. Minn, and L. Tamil, "ApneamedAssist: Real-time sleep apnea monitor using single-lead ECG," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp. 416-427, 2011.

[35] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc Natl Acad Sci USA*, vol. 88, pp. 2297-2301, 1991.

[36] G. Y. Bin, M. G. Shao, G. H. Bin, et al, "Detection of atrial fibrillation using decision tree ensemble," *in 2017 Computing in Cardiology*, 2017.

[37] S. Sarkar, D. Ritscher, and R. Mehra, "A detector for a chronic implantable atrial tachyarrhythmia monitor," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1219-1224, 2008.

A Robotic Sensor Node for Mechanical Property Detection of Material on Asteroid Surface

Jun Zhang*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
j.zhang@seu.edu.cn

Maozeng Zhang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
354926364@qq.com

Yizhuang Ding
School of Instrument Science and Engineering
Southeast University
Nanjing, China
220203519@seu.edu.cn

Liuchen Chen
School of Instrument Science and Engineering
Southeast University
Nanjing, China
220203518@seu.edu.cn

Qixuan Li
School of Instrument Science and Engineering
Southeast University
Nanjing, China
1017454559@qq.com

Minghan Qin
School of Instrument Science and Engineering
Southeast University
Nanjing, China
550749788@qq.com

Abstract—In recent years, asteroid exploration has attracted increasing attention from research institutes. One of the main tasks in the exploration is to acquire the material composition and mechanical property on the surface of the planets. This paper proposes an impacting based material composition and mechanical property detection method with a new kind of robotic sensor nodes. The sensor nodes have a cylindrical body and a conical head. During landing, the spacecraft launches the sensor nodes with a high speed to the surface of the asteroid. The sensor nodes impact the surface and record the vibration information with an accelerometer. We also used the sample entropy and detrended fluctuation analysis algorithms to obtain two characteristics of the signal. Then, the cluster analysis method was employed for vibration signal processing to determine the material composition and mechanical property on the surface of asteroids. We conducted experiments to validate the proposed method. The results of this paper can provide guidance for design of a detector for assisting of landing, anchoring, and sampling on asteroids to reduce the risks and increase the probability of success in these operations.

Keywords—Asteroid, exploration, robotic sensor node, mechanical property detection, vibration detection, signal processing

I. INTRODUCTION

With the development of deep space exploration technologies, asteroid exploration has gradually become a hotspot in recent decades [1]. Asteroids are small objects orbiting the Sun. They are thought to contain material from ancient times and are important for studying the origin of the solar system. Asteroids are also rich in mineral resources. For example, the asteroid Psyche contains 500 million tons of iron, 50 million tons of nickel, and hundreds of millions of tons of other rare metals as estimated [2]. It is foreseeable that with the maturity of technologies such as large-thrust rockets and sampling returns, asteroids can be used as the next destination for humans to mine mineral resources [3]. Some asteroids and comets even contain water. These asteroids can be used as a supply station for astronauts in deep space exploration [4]. Moreover, near-Earth asteroids have the risk of hitting the Earth. Studying the mechanism of asteroid orbital evolution and planetary defense strategy are crucial to human civilization [5-7].

So far, nearly 20 missions related to small celestial bodies have been launched at home and abroad [8], either flying over, or orbiting, or landing on the small bodies, or sampling as well as returning. All of them aimed to carry out scientific exploration of small celestial bodies in different forms. In the subsequent deep space exploration planning, NASA and JAXA actively promoted the implementation of small object sampling return tasks, mainly including the OSIRIS-REx [9], Hyabusa I [10], and Hyabusa II [11] asteroid sampling return missions.

The surface matter of asteroids mainly includes regolith, soil, pebbles, boulders, rock, and craters as reported [12]. The shape and surface material of the explored comet and asteroids are quite different [13-17]. There are many challenges in asteroid exploration because of the small gravity field [18] and unwell-known material composition, distribution, and mechanical property [19] which are quite different from the ones on Mars and Moon. Hence, landing and sampling on asteroids are extremely challenging and cannot take the same methods and strategies used on Mars and Moon.

The landing process includes several stages: lander bypass detection, close-range detection, landing point selection, and descent process planning [20]. In the landing process the surface topographic feature, material composition, distribution, and mechanical property should be detected. In the bypass detection stage, the lander uses radar or laser rangefinder to scan the asteroid appearance. The scanning data is processed by the software, and the three-dimensional model of the asteroid is reconstructed [21-22]. The rough data of the surface terrain can be obtained from the 3D model. Based on these data, several areas of interest should be selected [23]. The areas should be flatter, with no large raised hills or cliff gullies around them. Then the lander descents a certain height close to the observation area of interest. By using radar scanning, the lander draws a high-precision three-dimensional topographic map [24] and detects the composition of the asteroid surface, such as the distribution of the weathering layer, rock layer, and gravel heap. The lander also could employ the imaging ladar sensor sensors to sense the large rocks, meteorite pits, etc. on the surface and obtain the key geomorphological features [25]. Based on the detection of the topography, material composition, and distribution in several areas of interest, the lander synthesizes radar and image data to select a landing site [26].
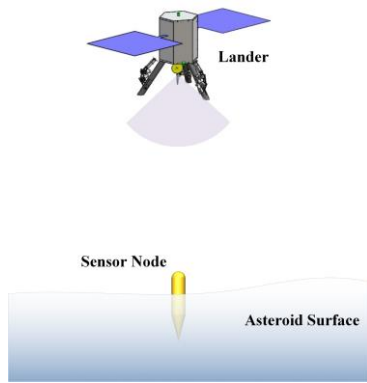
Fig. 1. Scenario of the robotic sensor node based material mechanical property detection of asteroid surface.

Fig. 2. Conditions showing different acceleration signals obtained from the impacts between the sensor node and the different samples.

After soft landing on the small bodies, anchoring [27-28] usually should be performed to fix the lander on the planet before sampling and other tasks. The material composition and mechanical property on the planet's surface decide the degree of difficulty of anchoring and sampling, the anchoring force and sampling efficiency. Hence, before anchoring and sampling, the material composition and mechanical property on the surface of asteroids also should be detected for preparing the anchoring and sampling tools and operation sequences. Detecting the mechanical properties of the asteroid surface provides strong support for the landing of the spacecraft on asteroids. At the same time, the material composition and mechanical property can also give information for the design and control of the mobile robots on the surface of asteroids [29-30].

Although the radar has been used to detect the composition and distribution of the surface materials in the planning of asteroid probe landing process and the regolith of Moon [31], it cannot detect the mechanical properties such as compressive strength and friction properties. Lacking these information will pose difficulties in contact buffering of lander landings and anchoring and sampling strategies planning [32-33]. Except for the usually used scientific instruments like radar, imaging ladar, laser rangefinder, near-infrared spectroscopy, impacting tools are simple and efficient devices for material composition and mechanical property detection of the asteroid surface. J. Biele et al. [34] reported the mechanical properties including strength and layering of two sites of the comet 67P/Churyumov-Gerasimenko based on the data recorded in the bounding and multiple landings of Philae. N. I. Kömle et al. designed a tethered harpoon as an anchoring device for Philae to detect the surface mechanical with an accelerometer sensor [35]. Unfortunately, the harpoon failed to fire in the mission [36]. K. Seweryn et al. also developed a low-velocity penetrometer for detecting the geotechnical properties of the planetary regolith [37]. Y. Ling et al. presented a method for rock identification with the sampling head rock during regolith sampling [38].

In this paper, an asteroid surface material composition detection method is proposed based on impacting with a new kind of robotic sensor nodes. The nodes can be shot into the asteroid surface to form a sensing network for a large area detection before landing, anchoring, and sampling. The system is overviewed in Section II. The detection method is talked about in Section III. The sensor node design and experimental validation are given in Section IV and V, respectively.
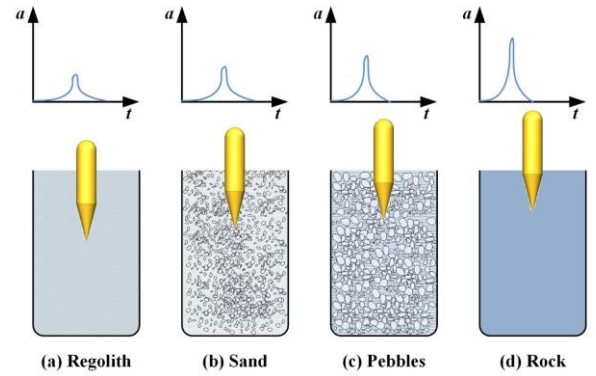
II. SYSTEM OVERVIEW

As shown in Fig. 1, in this paper, a scheme based on collision perception between small robotic sensor nodes and asteroids is proposed to detect the composition of substances and identify their mechanical properties. In the landing process, the lander senses the thickness of the weathering layer and selects the landing site with the image sensor. After selecting a proper site, the lander launches the sensor nodes with a given speed to the surface of the asteroid. The speed could be decided by mathematical modeling and experimental tests. The sensor nodes will hit the surface. The impact will create vibration signals. The sensor nodes can analyze the signal and classify the type of the material and estimate its mechanical properties. The sensor nodes form a wireless sensor network and collect the material information of a large area.

The sensor node impacting based method can provide future information for landing site determination. The method also can be used to determine the anchoring strategy before landing. This is very important because there is little time for anchoring when the lander touches on the microgravity surface of the asteroid. If the mechanical property of the asteroid is known well, the sampling method and sampling tools selection will be easier.

In this research, the impacting based perception mechanism needs to be studied. The relationship between the vibration under the conditions of different material composition with different mechanical properties, and different emission velocities should be investigated. On this basis, the sensor node will be designed to validate the proposed method.

III. METHOD

The composition of the surface material is obtained by shooting several small sensor nodes at a certain speed against the surface of the asteroid. The material mechanical properties detection is based on the sensor nodes. The vibration signal collecting by the sensor nodes is the first step. As shown in Fig. 2, we launch the sensor node into different materials, carry out multiple groups of measurements of the same material. Then, we analyze the data obtained and process the data of different materials to judge the composition of the samples. The accuracy and robustness of the scheme are also verified by analyzing the experimental data. This section mainly introduces the signal processing method used for the property determination and classification of the material.
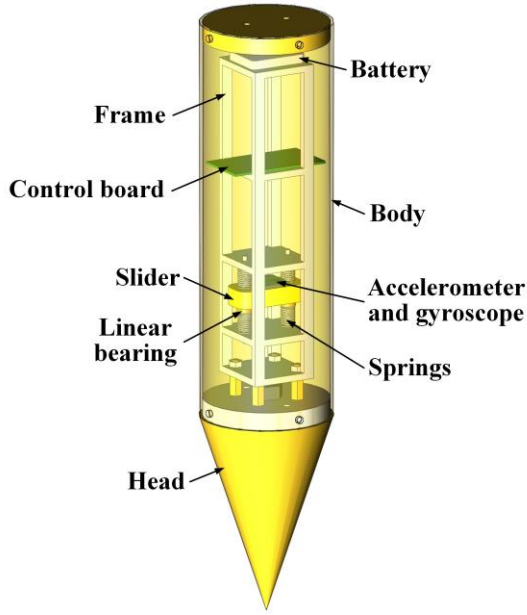
Fig. 3. Mechanism and components of the robotic sensor node.

The collision signal is a short data signal. To better handle the noisy short data signals common in physiological signals, Pincus proposed the algorithm of approximate entropy (ApEn). From a statistical point of view, the algorithm continues to maintain a similar conditional probability when the vector in the reconstruction space of the system increases from m dimension to $m+1$ dimension, reflecting the probability of a new pattern appearing in a time series during its evolution. And then measure the complexity of the time series. Therefore, ApEn algorithm can process short-term signals and has strong anti-interference ability, which is suitable for mixed signals composed of deterministic components and random components.

Although ApEn has many advantages, its calculation result is a biased estimate and lacks good relative consistency due to the self-matching included in the calculation process. To solve this shortcoming, Richman and Moorman developed a new algorithm named sample entropy (SampEn) which has better consistency than ApEn. The main steps of the SampEn algorithm are as follows:

Step 1: Assuming that the time series is $X = \{x_1, x_2, \cdots, x_N\}$, and then take m consecutive points to form a vector $X_m(i) = [x_i, x_{i+1}, \cdots, x_{i+m-1}]$, $i = 1, 2, \cdots, N-m$. There are $N-m$ m-dimensional vectors in total.

Step 2: Define the distance between vectors $X_m(i)$ and $X_m(j)$ as $d[X_m(i), X_m(j)] = \max(|x_{i+k}-x_{j+k}|)$, where $k=0, 1, \cdots, m-1$, $i, j = 1, 2, \cdots, N-m$, $i \neq j$.

Step 3: Set the threshold to r, and count the number of elements in the $N-m$ vectors that do not include $X_m(i)$, whose distance to the element $X_m(i)$ is less than the threshold, and record the number as the template matching number $N_m(i)$. Let $C_m^r(i) = N_m(i)/(N-m-1)$. The average of $C_m^r(i)$ is:

$$\Phi_m(r) = \frac{\sum_{i=1}^{N-m} C_m^r(i)}{N-m} \qquad (1)$$

Step 4: Increase the dimension of the space to $m + 1$, and calculate $\Phi_{m+1}(r)$ according to the above steps.

Step 5: The SampEn of the time series X is:

$$\text{SampEn}(m, r) = \lim_{N \to \infty} \ln \Phi_m(r)/\Phi_{m+1}(r) \qquad (2)$$

when N is a finite number, SampEn is calculated as:

$$\text{SampEn}(m, r, N) = \ln \Phi_m(r)/\Phi_{m+1}(r) \qquad (3)$$

The algorithm is used to calculate the SampEn of acceleration data collected from different kinds of materials experimental tests.

Except for the SampEn algorithm, the Detrended Fluctuation Analysis (DFA) algorithm is also selected to obtain the complexity and regularity of the materials. The main steps of the algorithm are as follows:

Step 1: Assuming that the time series is $X = \{x_1, x_2, \cdots, x_N\}$, then we can obtain the integrated time series as

$$y(i) = \sum_{j=1}^{i} [x(j) - \bar{x}] \qquad (4)$$

where $\bar{x}$ is the mean of the time series X.

Step 2: Divide the time series $y(i)$ into intervals of length s without overlapping. The number of the intervals is $B=[N/s]$.

Step 3: Detrend all the intervals

$$y_s(i) = y(i) - p_m(i) \qquad (5)$$

Where $p_m(i)$ is the n-order fitted polynomial of interval m. $n = 1, 2, 3, \cdots$.

Step 4: Calculate the mean square error of each interval after filtering the trend

$$F_s^2(m) = y_s^2(i) = \frac{1}{s} \sum_{i=1}^{s} y_s^2[(m-1)s + i] \qquad (6)$$

Step 5: Calculate the square roots of B detrending subintervals to obtain the DFA fluctuation function

$$F(s) = \sqrt{\frac{1}{B} \sum_{m=1}^{B} F_s^2(m)} \qquad (7)$$

If the logarithmic curve of F(s) and s has a linear relationship

$$\lg(F(s)) = \lg(c) + \alpha \lg(s) \qquad (8)$$

then, we can obtain the power-law wave equation

$$F(s) = cs^\alpha \qquad (9)$$

The slope α (DFA scale index) of the straight line can be calculated by the least square method. The difference in the scale index α reflects the correlation of the time series.

Based on the two algorithms above, we can obtain two characteristics, the SampEn and scale index α. Then, the Cluster Analysis method can be used for classifying the surface materials with different compositions.

IV. ROBOTIC SENSOR NODE DESIGN

A. Mechanism

To verify the feasibility of the impacting based method, we have designed a simplified version of the sensor node, which is made up of a hollow cylinder body and a conical head as shown in Fig. 3. The vibration signal recording mechanism and components are inside the body installing on the frame. In

Fig. 4. Prototype of the robotic sensor node.

Fig. 5. Experimental setup for the material mechanical property detection.

order to avoid damage to the sensor module, we designed a simple buffer mechanism which consists of two slide ways, two linear bearings, a slider, and four springs. The inertial sensor unit is mounted on the upper side of the slider. The mechanism can reduce the impact acceleration of the sensor unit when the sensor node hits the too hard rocky surface of asteroids. The control board and battery are installed on the upper floors of the frame.

B. Hardware

We use the WT931 module produced by wit-motion as the acceleration sensing component of the sensor node. This module integrates a high-precision gyroscope, accelerometer, and geomagnetic sensor, and adopts a high-performance microprocessor and algorithm to process the data of the sensor node. The acquisition frequency of the module can be set up to 500Hz. In this work, we mainly use the data from the three-axis accelerometer. The data collected by the accelerometer is passed to an STM32F103 MCU via the I2C bus. Then the MUC saves the data to the 16GB SD card manufactured by Kingston Company through the RS232 interface. The SD card is placed in the Openlog card slot and can be removed from the sensor node at the end of the experiment, and the data is imported into the computer through the USB3.0 card reader. The sensors and MCU are powered by a 3.7V polymer lithium battery with the capacity of 1000mAh.

C. Software

The software system of the node is mainly programmed and designed by Keil uVision5 and downloaded to our MCU. The function of the software is to read the original hexadecimal data of JY901. Then the data is transmitted to the Openlog module by the MCU. The Openlog module holds the data in .txt form of the original acceleration data. After that, the .txt file is read and processed with a PC, and the triaxial acceleration information can be obtained. Finally, the acceleration of each impacting motion is processed by MATLAB, and the influence of different materials on the acceleration of the nodes is analyzed.

V. Experiments

A. Prototype and Test Platform

A prototype of the sensor node was implemented for experimental studies. We made a 304 steel cylinder body with a diameter of 60 mm and a length of 180mm. A top cover and a bottom cover were installed at both ends of the body. A
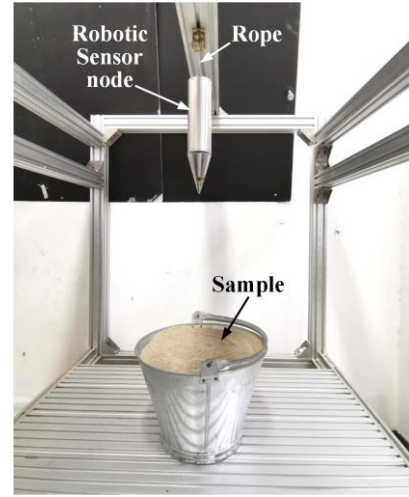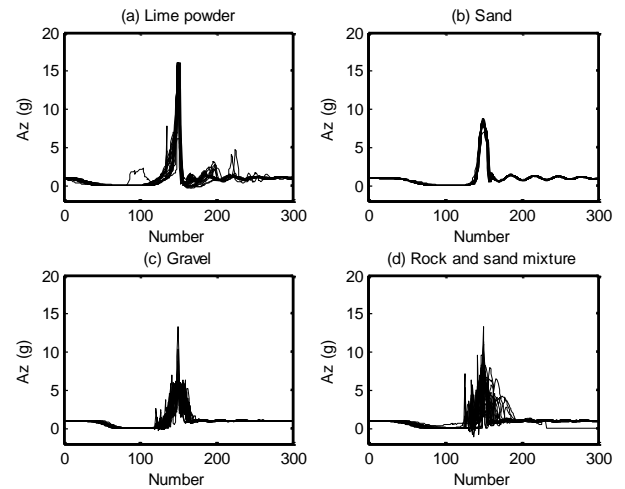
Fig. 6. The change of the accelerations during the impacts of the three kinds of materials.

stainless steel cone with a bottom diameter of 60mm and a height of 103mm was also designed attached to the bottom cover to simulate the head of the sensor node. The internal circuit of the sensor node is mounted on an L-shaped metal plate and is fastened to the inside of the cylinder by screws. The prototype of the sensor node is shown in Fig. 4 with the total height of 28 cm and mass of about 1.5 kg.

As depicted in Fig. 5, in order to simulate the scene of the sensor node penetrating the asteroid surface, we designed a platform for experimental tests. The sensor node was fixed on the platform with a rope above a sample bucket. Different kinds of sample simulants were contained in the bucket to simulate different experimental materials. When the experiment was carried out, we cut the hanging rope of the sensor node. The sensor node fell into the bucket and eventually stopped moving under the action of the simulants.

B. Different Sample Simulant Detection Tests

We used four kinds of substances i.e. lime powder, sand, gravel, and rock-sand mixture as sample simulants in the experimental tests. In the experiments, the sensor nodes fell freely from the height of 100mm above the surface of the simulants. Each sample simulant measurement was repeated 30 times. The collected acceleration data were used for
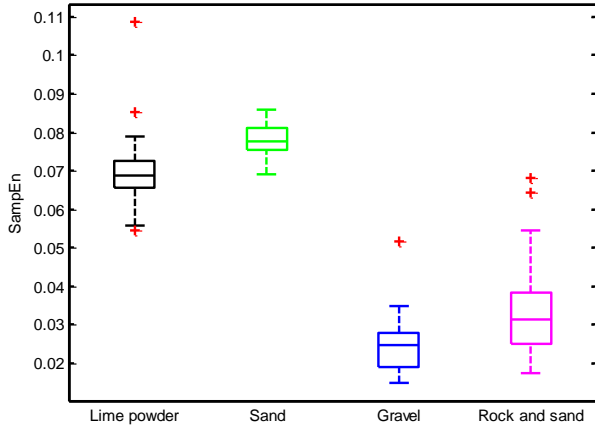
Fig. 7. The distribution of the entropy value for the impacting data of the different kinds of materials.

Fig. 8. The distribution of the scale index α for the impacting data of the different kinds of materials.

classifying. The data recording frequency was set as 500Hz. We took the acceleration value of the impacting direction in each set of the measurement data for processing by MATLAB.

C. Vibration Signal Processing and Sample Classifying

In order to compare the collected acceleration data easily, we find the maximum of each recorded data set to align the data sets. The 300 data points of each data set were selected for the processing and classification. The change of the acceleration in the measurements is shown in Fig. 6. The results show that the lime powder has the largest acceleration change than the other materials. This was caused by the too loose property of the lime powder. The sensor node touched the bottom of the sample container as we found in the experiments. The sand produced acceleration signals with good consistency and had the smallest extremum but the longest decay time. The gravel also caused large acceleration which had large fluctuations but short decay time. The rock and sand mixture had the most complicated changes in the four kinds of materials and the largest fluctuations which showing the characteristics of the mixture. The maximum acceleration and the decay time could be used to identify the material composition roughly.

We also used sample entropy to analyze the entropy value distribution of the acceleration data. Fig. 7 shows the entropy results of the four kinds of sample simulants. It can be seen from the figure that the entropy value of the lime powder and sand are significantly higher than that of the other two materials. This shows that the time series generated by impacting gravel and rock-sand mixture have a low degree of disorder, while the time series generated by impacting lime powder and sand have a high degree of disorder.

Fig. 8 shows the DFA analysis results. It can be seen from the distribution of α that the impact time series under the four simulants have a non-power-law long-range correlation. Based on the two characteristic values of sample entropy and DFA scale factor α, cluster analysis was conducted on the data sets. The number of iterations was set as 15 times. The clustering results are shown in Fig. 9. The four materials correspond to 120 recorded time series.

The clustering results indicate that Type II and III are relatively similar, while Type I and IV are significantly different from the other two types. Types I, II, III, and IV correspond to lime powder, sand, gravel, and rock-sand
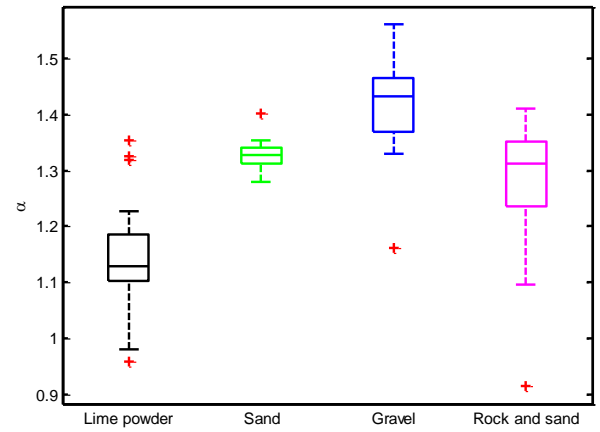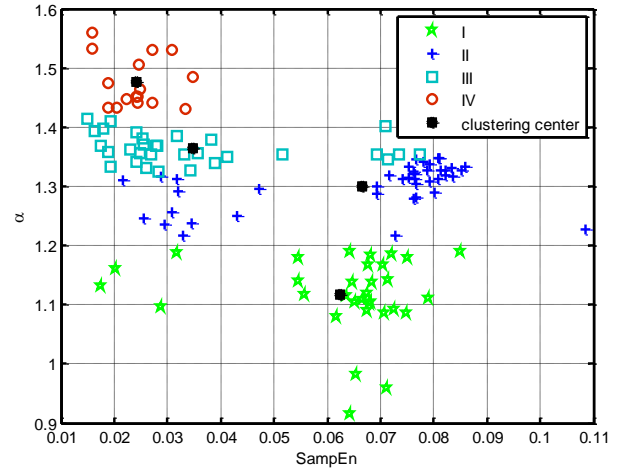
Fig. 9. The clustering results of the impacting data of the different kinds of materials.

mixture, respectively. It can be judged that the impact data under these four materials can be distinguished.

VI. CONCLUSIONS AND FUTURE WORK

This paper proposes a detection method based on sensor nodes to sense the surface of asteroids. Several small sensor nodes are driven into the surface of the planet by the asteroid lander, and the measured node acceleration is fed back to the control processing unit to determine the mechanical properties of the asteroid surface material. This determines the composition of the surface material of the asteroid. The method has the advantage of low complexity and high stability. In this paper, the effectiveness of this scheme is verified by experiments. It is concluded that gravel is the most effective material for the deceleration of the sensor node, which is in line with our expectations. The four materials were classified by processing their acceleration signals with the sample entropy and detrended fluctuation analysis algorithms, and the cluster analysis method.

Future work will focus on designing landing simulation platforms for field experiments. More kinds of material will be used in the tests to obtain larger data sets for material classification. The compressive hardness, angle of internal friction, cohesion, etc. on the influences of the acceleration signal will be investigated quantitatively to classify the material more precisely.

REFERENCES

[1] P. Vere, R. Jedicke, R. Wainscoat, "Detection of earth-impacting asteroids with the next generation all-sky surveys," *Icarus*, vol. 203, no. 2, pp. 472-485, 2009.

[2] L. T. Elkins-tanton, J. Bell, et al., "Asteroid (16) psyche: the science of visiting a metal world," *Lunar & Planetary Science Conf.*, 2016.

[3] M. J. Sonter, "Mineral resource estimation for asteroid mining projects," *Lunar and Planetary Science Conf.*, 2014.

[4] V. Badescu, "Asteroids: prospective energy and material resources," Springer, 2013, DOI10.1007/978-3-642-39244-3.

[5] Chapman, Clark R., and D. Morrison, "Impacts on the earth by asteroids and comets: assessing the hazard," *Nature*, vol. 367, no. 6458, pp. 33-40, 1994.

[6] K. Kosmo, P, Lubin, G. B. Hughes, J. Griswold, Q. Zhang, et al., "Directed energy planetary defense," in: *IEEE Aerospace Conf.*, Big Sky, MT, USA, 7-14 March 2015, pp. 1-9.

[7] P. Lubin, G. B. Hughes, M. Eskenazic, K. Kosmo, I. E. Johansson, et al., "Directed energy missions for planetary defense," *Adv. Space Res.*, vol. 58, no. 6, pp. 1093-1116, Sep. 2016.

[8] P. N. Desai, R. A. Mitcheltree, F. M. Cheatwood, "Sample Returns Missions in the Coming Decade," *NASA Langley Technical Report Server*, 2000.

[9] K. Sankaran, B. Hamming, C. Grochowski, et al., "Evaluation of existing electric propulsion systems for the OSIRIS-REx mission," *J. Spacecr. Rocket.*, vol. 50, no. 6, pp. 1292-1295, 2013.

[10] M.A. Shoemaker, J.C.V.D. Ha, S. Abe, et al., "Trajectory estimation of the Hayabusa spacecraft during atmospheric disintegration," *J. Spacecr. Rocket.*, vol. 50, no. 2, pp. 326-336, 2013.

[11] Y. Tsuda, M. Yoshikawa, M. Abe, et al., "System design of the Hayabusa 2-Asteroid sample return mission to 1999 JU3," *Acta Astronaut.*, vol. 91, pp. 356-362, 2013.

[12] L. A. M. Benner, S. J. Ostro, C. Magri, et al., "Near-Earth asteroid surface roughness depends on compositional class," *Icarus*, vol. 198, no. 2, pp. 294-304, 2008.

[13] A. Coradini, F. Capaccioni, S. Erard, G. Arnold, M. C. De Sanctis, "The Surface Composition and Temperature of Asteroid 21 Lutetia As Observed by Rosetta/VIRTIS," *Science*, vol. 334, no. 6055, pp. 492-494, Oct. 2011.

[14] N. Thomas, H. Sierks, C. Barbieri, P. L. Lamy, R. Rodrigo, et al., "The morphological diversity of comet 67P/Churyumov-Gerasimenko," *Science*, vol. 347, no. 6220, pp. aaa0440, Jan. 2015.

[15] K. Kitazato, R. E. Milliken, T. Iwata, M. Abe, M. Ohtake, "The surface composition of asteroid 162173 Ryugu from Hayabusa2 near-infrared spectroscopy," *Science*, vol. 364, no. 6437, pp. 272-275, Apr. 2019.

[16] S. Sugita, R. Honda, T. Morota, S. Kameda, H. Sawada, "The geomorphology, color, and thermal properties of Ryugu: Implications for parent-body processes," *Science*, vol. 364, no. 6437, pp. eaaw0422, Apr. 2019.

[17] D. S. Lauretta, D. N. DellaGiustina, C. A. Bennett, D. R. Golish, K. J. Becker, et al., "The unexpected surface of asteroid (101955) Bennu," *Nature*, vol. 568, pp. 55–60, Mar. 2019.

[18] Y. Takahashi, D. J. Scheeres, and R. A. Werner, "Surface gravity fields for asteroids and comets," *J. Guid. Control. Dyn.*, vol. 36, no. 2, pp. 362-374, 2013.

[19] S. A. Sipila, Z. C. Scoville, J. T. Bowie, and J. A. Buffington, "Extravehicular activity asteroid exploration and sample collection capability," in: *Int. Conf. on Space Oper.*, 2014, pp. 1-11.

[20] B. Leroy, "Crater detection for autonomous landing on asteroids," *Image Vision Comput.*, vol. 19, no. 11, pp. 787-792, 2001.

[21] R. A. Werner and D. J. Scheeres, "Exterior gravitation of a polyhedron derived and compared with harmonic and mascon gravitation representations of asteroid 4769 Castalia," *Celest. Mech. Dyn. Astr.*, vol. 65, no.3, pp. 313-344, 1996.

[22] M. T. Zuber, D. E. Smith, A. F. Cheng, J. B. Garvin, O. Aharonson, et al., "The Shape of 433 Eros from the NEAR-Shoemaker Laser Rangefinder," *Science*, vol. 289, no. 5487, pp. 2097-2101, Sep. 2000.

[23] R. Furfaro, D. Cersosimo, and D. R. Wibben, "Asteroid precision landing via multiple sliding surfaces guidance techniques," *J. Guid. Control Dynam.*, vol. 36, no.4, pp.1075-1092, 2013.

[24] T. D. Cole, M. T. Zuber, G. Neuman, A. F. Cheng, R. A. Reiter, et al., "Analysis of laser radar measurements of the asteroid 433 Eros," in: *Proc SPIE Int Soc Opt Eng*, 4377, pp. 163-174, 2001.

[25] Wheel, Peter J, M. E. Dobbs, and W. E. Sharp, "Optimization of space borne imaging ladar sensor for asteroid studies using parameter design," in: *Proc SPIE Int Soc Opt Eng*, pp. 68-77, 2002.

[26] J. Li, J. Zhen and X. Ruan, "Optimal site selection for soft landing on asteroid," in: *Chin. Control Decis. Conf.*, Changsha, 2014, pp. 2944-2948.

[27] J. Zhang, C. Dong, A. Song, "Force modeling of the cutting disc in rock sawing for anchoring and sampling in asteroid exploration," in: *IEEE Int. Conf. Robot. Biomim*, Macau, China, Dec. 2017, pp. 2142-2147.

[28] J. Zhang, C. Dong, H. Zhang, S. Li, A. Song, "Modeling and experimental validation of cutting based lander anchoring and sampling methods for asteroid exploration," *Adv. Space Res.*, vol. 61, no. 9, pp. 2426-2443, May 2018.

[29] D. Helmick, B. Douillard, and M. Bajracharya, "Small body surface mobility with a limbed robot," in: *IEEE/RSJ Int. Conf. Intell. Rob. Syst.*, Chicago, IL, USA, 14-18 Sept. 2014, pp. 2341-2348.

[30] Y. Yuguchi, W. F. R. Ribeiro, K. Nagaoka, and K. Yoshida, "Experimental evaluation of gripping characteristics based on frictional theory for ground grip locomotive robot on an asteroid," in: *IEEE Int. Conf. Rob. Autom.*, Seattle, WA, USA, 26-30 May 2015, pp. 2822-2827.

[31] J. Zhang, W. Yang, S. Hu, Y. Lin, G. Fang, et al., "Volcanic history of the Imbrium basin: A close-up view from the lunar rover Yutu," *PNAS*, vol. 112, no. 17, pp. 5342-5347, Apr. 2015.

[32] K. Zacny, P. Chu, G. Paulsen, et al., "Asteroids: anchoring and Sample acquisition approaches in support of science, exploration, and in situ resource utilization," *Asteroids*, Springer, Berlin, Heidelberg, pp 287-343, 2013.

[33] S. Ulamec and J. Biele, "Surface elements and landing strategies for small bodies missions – Philae and beyond," *Adv. Space Res.*, vol. 44, no.7, pp. 847-858, Oct. 2009.

[34] J. Biele, S. Ulamec, M. Maibaum, R. Roll, L. Witte, et al., "The landing(s) of Philae and inferences about comet surface mechanical properties," *Science*, vol. 349, no. 6247, pp. aaa9816, Jul. 2015.

[35] N. I. Kömle, A. J. Ball, G. Kargl, J. Stöcker, M. Thiel, et al., "Using the anchoring device of a comet lander to determine surface mechanical properties," *Planet. Space. Sci.*, vol. 45, no. 12, pp. 1515-1538, Dec. 1997.

[36] E. Hand, "Philae probe makes bumpy touchdown on a comet," *Science*, vol. 346, no. 6212, pp. 900-901, 2014.

[37] K. Seweryn, K. Skocki, M. Banaszkiewicz, J. Grygorczuk, M. Kolano, et al., "Determining the geotechnical properties of planetary regolith using Low Velocity Penetrometers," *Planet. Space. Sci.*, vol. 99, pp. 70–83, Sep. 2014.

[38] Y. Ling, W. Lu, and A. Song, "Sampling head-rock contact identification for regolith sampling in space," *Aerosp. Sci. Technol.*, vol. 31, no. 1, pp. 108-114, 2013.

Research and Verification of Multi-Satellite Thermal Vacuum Test Method

Xiaofeng Zhang*
Scientific Satellites Division
Innovation Academy for Microsatellites
of CAS
Shanghai, China
Zhangxf@microsate.com

Li Wu
Telecommunication Satellites Division
Innovation Academy for Microsatellites
of CAS
Shanghai, China
Wul@microsate.com

Hong Liu
Telecommunication Satellites Division
Innovation Academy for Microsatellites
of CAS
Shanghai, China
Liuh@microsate.com

Jianchao Feng
Scientific Satellites Division
Innovation Academy for Microsatellites
of CAS
Shanghai, China
Fengjc@microsate.com

Meijuan Xu
AIT center
Innovation Academy for Microsatellites
of CAS
Shanghai, China
Xumj@microsate.com

Rui Cheng
project management Department
Innovation Academy for Microsatellites
of CAS
Shanghai, China
Chengr@microsate.com

Abstract—**A method for simultaneous test with multi-satellite was presented in this paper for the problem of production and high-efficiency thermal vacuum test with multi-satellite. Which was applied in the vacuum thermal test of a certain type of satellites after theoretical analysis and numerical simulation. This method was proved to be effective and feasible by the test and on-orbit results. Simultaneous test with multi-satellite saved test time in contrast with that with single satellite, which also reduced test and labor costs and improved the cost-effectiveness.**

Keywords—batch satellites; thermal vacuum test; simultaneous test with multi-satellite; cost-effectiveness

I. INTRODUCTION

In order to verify the overall function and performance of the spacecraft in space environment and ensure the reliable operation of the spacecraft in orbit, it is necessary to carry out sufficient vacuum thermal test (including thermal balance test and thermal vacuum test) on the ground. Among them, the thermal balance test is mainly to verify the correctness of thermal design, assess the capability of thermal control subsystem, obtain the temperature data of the whole spacecraft and update the mathematical model of thermal analysis; The thermal vacuum test is for each satellite to be launched, to expose the defects of the satellite in design, material and manufacturing process, eliminate early failures, and evaluate the performance of the entire satellite [1]. Spacecraft vacuum thermal test is the most complicated, cost-intensive, and time-consuming test project in the development process. It is also an effective and necessary means to improve the reliability of spacecraft in orbit [2].

In recent years, the microsatellite industry has grown stronger, and a large number of resources including many commercial companies which have entered the field, and plan to establish various microsatellite constellation projects. Google's "O3b Networks" plan that includes 120 satellites constellation has completed 3 launches of 12-satellite satellites. SpaceX successfully launched six OrbCOMM-G2 communications satellites in July 2014, followed by Elon Musk's 400-satellite Constellation plan to build a global high-speed Internet; In August 2015, Samsung also proposed the "Air Internet" plan, aiming to launch 4600 low-orbit microsatellites and build a global low-cost air network [3]. It can be seen that the satellite industry is developing towards mass and miniaturization. Meanwhile, all countries in the world are committed to reducing the development cost of space projects [4], among which parallel test is one of the effective methods to reduce the development cost. Beijing Satellite Environmental Research Institute once completed the thermal vacuum environmental strain test of two satellite antenna arrays in the KM4 space simulator [5].

The traditional single-satellite thermal test method of spacecraft has a long development cycle, high cost and low efficiency. In order to adapt to the rapid, flexible and low-cost development and production requirements of current batch satellites projects, this paper puts forward the method of multi-satellite thermal test at the same time. Compared with the traditional single satellite thermal test in a vacuum chamber, multi-satellite thermal test in a vacuum chamber at the same time will encounter the following problems and challenges:

- The consistency of the temperature field at the high and low temperature end of multiple satellites tested at the same time. Due to the mutual thermal interference between the satellites and the different positions of the satellites in the vacuum chamber, the thermal environment conditions of the satellites are inconsistent, resulting in inconsistent temperature fields of the satellites. On the one hand, it is difficult to control the temperature, on the other hand, some stand-alone units may not meet the temperature index requirements of the thermal vacuum test;

- The problem of inconsistent heating and cooling rates. Due to the inconsistency of thermal environmental conditions between satellites, the rise and fall rates are inconsistent, which increases the difficulty of temperature control and extends the test time;

- The problem of test duration. Compared with single-satellite test, multi-satellite test increases the thermal load of vacuum simulator. On the other hand, mutual shielding reduces the view factor of the satellite against heat sink, resulting in a slower cooling rate. In addition, the control difficulty of the multi-satellite simultaneous test itself increases, thus extending the total test duration. If the cumulative total time of multi-satellite simultaneous test is not significantly reduced compared with the single test of multiple satellites, it unable to achieve the purpose of the multi-satellite

simultaneous test to improve the cost-effectiveness ratio.

Based on a batch satellite project, the feasibility analysis of multi-satellite simultaneous test was carried out in this paper. Subsequently, the parallel thermal vacuum test of three satellites entering the vacuum chamber at the same time was completed. The test results met the requirements of the mission and the on-orbit functions and performance of the satellite during its lifetime were normal.

II. FEASIBILITY STUDY OF MULTI-SATELLITE TEST

A. The Theoretical Analysis

In order to explore the influence factors of satellite temperature in vacuum chamber, the theoretical analysis is carried out. The theoretical basis of thermal test is the radiation heat transfer system of two closed cavities composed of two diffuse gray surfaces as shown in Fig.1 [6], where A is a fully convex surface, similar to the satellite in the vacuum chamber, and B is a closed cavity, similar to the vacuum chamber.
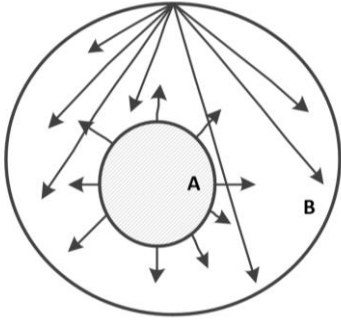
Fig.1.   A radiative heat transfer system consisting of two objects

Heat transfer $\emptyset_{A,B}$ between A and B can be written as:

$$\emptyset_{A,B} = \frac{A_A(E_{bA} - E_{bB})}{\left(\frac{1}{\varepsilon_A} - 1\right) + \frac{1}{X_{A,B}} + \frac{A_A}{A_B}\left(\frac{1}{\varepsilon_B} - 1\right)}$$

$$E_b = \sigma T^4$$

In the formula: A_A is the surface area of A; E_{bA} and E_{bB} are the black body radiation intensity of A and B; ε_A and ε_B are the emissivity of A and B surface; $X_{A,B}$ is the radiant interchange factor of surface A to B. It can be seen that the satellite surface temperature as follows:

$$T_A = f(T_B, A_A, A_B, \varepsilon_A, \varepsilon_B, X_{A,B}).$$

Taking three satellites as an example, physical modeling and calculation of three-satellite simultaneous test are carried out. The positions of the three satellites in the tank are shown in Fig.2. The numbers 1, 2, and 3 are the satellite numbers. The satellites are simplified to a 1m regular hexahedron with a distance of 0.7m. The vacuum tank has a diameter of 3.5m and a depth of 5.5m. Simplify and assume as follows [7]:

- The internal heat source is the same, and the external heat flow is simulated by the infrared heating cage, which is an isothermal body with uniform surface temperature;

- To simplify the heat exchange between satellites, mainly considering the shielding of the heat sink between each other, that is, it is assumed that the temperature difference between the three satellites is not large, and the radiative heat exchange between each other is negligible relative to the heat exchange with the heat sink;
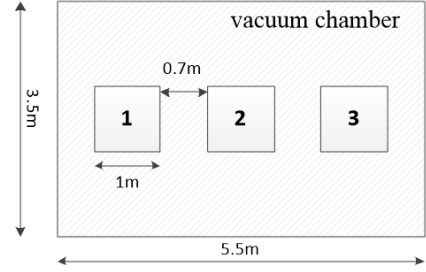
Fig.2.   The layout of three satellites in vacuum chamber

Due to symmetry, only the differences between satellite 1 and satellite 2 are analyzed. Referring to the radiative heat transfer system of a closed cavity composed of two dusty surfaces, the following heat transfer calculation formulas can be written for satellite 1 to vacuum tank (indicated by subscript KM) and satellite 2 to vacuum tank respectively:

$$\emptyset_{1,KM} = \frac{A_1(E_{b1} - E_{bKM})}{\left(\frac{1}{\varepsilon_1} - 1\right) + \frac{1}{X_{1,KM}} + \frac{A_1}{A_{KM}}\left(\frac{1}{\varepsilon_{KM}} - 1\right)}$$

$$\emptyset_{2,KM} = \frac{A_2(E_{b2} - E_{bKM})}{\left(\frac{1}{\varepsilon_2} - 1\right) + \frac{1}{X_{2,KM}} + \frac{A_2}{A_{KM}}\left(\frac{1}{\varepsilon_{KM}} - 1\right)}$$

In engineering, the calculation method of view factor of two parallel surfaces as shown in Fig.3 is as follows [6] :

$$X_{i,j} = \frac{\left[(W_i + W_j)^2 + 4\right]^{1/2} - \left[(W_j - W_i)^2 + 4\right]^{1/2}}{2W_i}$$

$$W_i = w_i/L$$
$$W_j = w_j/L$$

Fig.3.   Calculationg method of view factor

If the same external heat flux is considered, the temperature difference of each satellite in the thermal equilibrium state can be calculated, and let $\emptyset_{1,KM} = \emptyset_{s,KM}$, $\emptyset_{2,KM} = \emptyset_{s,KM}$, available:

$$\frac{T_1^4 - T_{KM}^4}{T_s^4 - T_{KM}^4} = \frac{1}{C_{1s}}; \quad \frac{T_2^4 - T_{KM}^4}{T_s^4 - T_{KM}^4} = \frac{1}{C_{2s}}$$

Generally, the surface temperature and surface emissivity of the vacuum chamber are fixed[8]. It can be seen from the

analysis that the larger the values of $X_{1,KM}$ and $X_{2,KM}$ are, the smaller the A_s/A_{KM} and the ε_h are, the larger the values of C_{1s} and C_{2s} are and the closer to 1. Therefore, in order to reduce the temperature difference between multi-satellite test and single-satellite test, we should try to increase the view factor between satellite and thermal sink, reduce the ratio of satellite surface area to thermal sink internal surface area A_s/A_{KM}, or reduce the ε_h of satellite surface emissivity.

According to the geometric size of the existing satellites, vacuum chambers of appropriate specifications and the number of satellites tested at the same time should be selected to make the surface area of the satellites and the internal surface area of the heat sink relatively small. At the same time, the distance between satellites should be as far as possible in the layout, and each other can be stagger to reduce the shielding effect, thus increasing the angular coefficient of the satellites to the heat sink.

The effect of satellite surface emissivity ε_h was analyzed as follows.

Fig.2 shows $X_{1,KM} = 0.913$, $X_{2,KM} = 0.826$, $A_s/A_{KM} \approx 0.075$ and $\varepsilon_{KM} = 0.9$ were substituted into the expression, and several typical satellite surface emissivities of 0.9 (black paint), 0.85 (white paint), 0.6 (polyimide film), and 0.03 (equivalent emissivity of MLI) were taken. Table I shows the variation.

TABLE I. THE VARIATION OF C_{s1} AND C_{s2} WITH ε_h CHANGES

ε_h	C_{s1}	C_{s2}	C_{s2}/C_{s1}
0.9	0.921545	0.841592	0.91324
0.85	0.925551	0.849012	0.917304
0.6	0.946166	0.888262	0.938801
0.03	0.99715	0.99372	0.996561

It can be seen from the data in the table that as ε_h decreases, the values of C_{s1} and C_{s2} increase significantly and are closer to 1, that is, the difference in heat exchange between each satellite during the multi-satellite test and its individual test is smaller . Since satellite 2 is blocked on both sides, compared with satellite 1, the heat flow between satellite 2 and a single satellite is quite different.

TABLE II. RELATIONSHIP BETWEEN CHANGES IN ε_h (T_s=305K) (TEMPERATURE UNIT: K)

ε_h	T_1	$T_1 - T_s$	T_2	$T_2 - T_s$
0.9	311.2	6.2	318.3	13.3
0.85	310.8	5.8	317.6	12.6
0.6	309.2	4.2	314.1	9.1
0.03	305.2	0.2	305.5	0.5

TABLE III. RELATIONSHIP BETWEEN CHANGES IN ε_h (T_s=265K) (TEMPERATURE UNIT: K)

ε_h	T_1	$T_1 - T_s$	T_2	$T_2 - T_s$
0.9	270.4	5.4	276.5	11.5
0.85	270.1	5.1	275.9	10.9
0.6	268.6	3.6	272.8	7.8
0.03	265.1	0.1	265.4	0.4

The value of C_{s2}/C_{s1} characterizes the difference in heat transfer between satellites during the multi-satellite test. Similarly, as ε_h decreases, the closer the value of C_{s2}/C_{s1} is

to 1, the difference in heat transfer between the satellites is smaller.

Based on the results in Table 1, the variation relationship of the temperature difference among satellites with the variation of ε_h was further obtained. Assuming that the heat sink temperature TKM=100K, Ts=305K and Ts=265K were taken to analyze the temperature difference between satellites at high and low temperatures. The calculation results are shown in Table II and Table III. The results show that under the same external heat flow conditions, the equilibrium temperature of satellite 1 and satellite 2 is higher than that of a single satellite, and as the surface emissivity of the satellite decreases, the temperature difference also decreases. When the surface of the satellite is fully covered with MLI(multi-layer insulation blanket), the temperature difference at the high and low temperature ends is less than 0.5°C. The temperature difference between the satellite 2 and the single satellite test is roughly twice that of the satellite 1 and the single satellite test. By comparing Table II and Table III, it can be seen that as the equilibrium temperature increases, the temperature difference between the multi-satellite test and the single-satellite test also increases slightly.

Therefore, the MLI should be coated in accordance with the satellite technical status during the test, and the coating should be as complete as possible to reduce heat leakage, which can reduce the temperature difference between satellites and that between a single satellite during the test.

Based on the above analysis and combined with the actual situation of the project, the scheme of three satellites entering KM3 vacuum chamber at the same time for thermal vacuum test was adopted. Fig. 4 shows the satellite layout in the chamber. Satellite 2 is staggered by a certain distance from satellite 1 and satellite 3 in the Y direction, thus increasing the angular coefficient of each satellite and heat sink. (Note: Due to interference with the satellite's extruded portion (not shown), satellite 2 was installed with a rotation of 180° around the Z-axis relative to satellite 1.)

Surface ±Y of the satellite is the main heat dissipation surface, while surface ±X of the satellite is mainly covered with MLI. During installation, surface ±Y is oriented towards heat sink, and surface ±X is opposite to each other, thus reducing thermal interference between satellites.
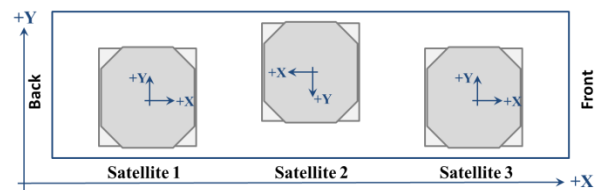
Fig.4. The actual layout of three satellites in vacuum chamber for thermal vacuum test

B. Simulation Analysis

In order to analyze the balance temperature difference and cooling rate difference of each satellite in the test scheme more accurately, the thermal simulation model based on a certain satellite model was used to carry out numerical analysis and verification of multi-satellite simultaneous test.

1) Simulation modeling

Fig. 5 shows the satellite is divided into nodes. The yellow area outside the satellite is the heat dissipation surface, the

light blue area is MLI blanket, the green area inside the satellite is the structure panel, and the dark blue is the stand-alone instrument. The coordinate system in the picture is the satellite body coordinate system of the No.2 satellite.
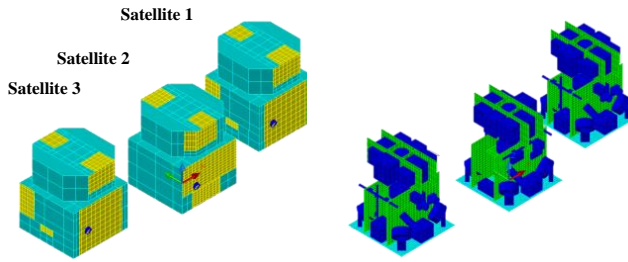
Fig.5. The geometry model of 3 satellites (left: the surface of 3 satellites, right: inside of 3 satellites)

Fig. 6 shows the layout in the KM3 vacuum chamber.
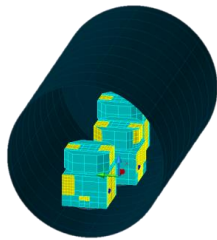
Fig.6. The layout of three satellites in vacuum chamber

The vacuum chamber was set as the constant temperature boundary condition (-180 ℃), the surface emissivity was 0.9[9], and the actual external heat flow of the satellite was given for simulation calculation. The temperature of three satellites and single satellite were calculated respectively.

2) Equilibrium temperature deviation analysis
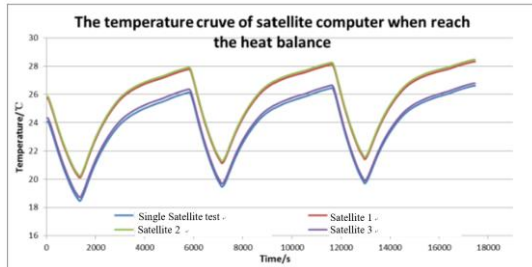
Fig.7. The temperature curve of onboard computer
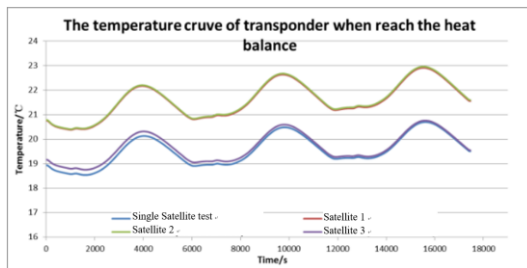
Fig.8. The temperature curve of TM&TC units

The external heat flow under a certain high temperature condition was selected for calculation, and the temperature curve of thermal balance of two typical instruments, the onboard computer and the TM&TC unites, was analyzed, as shown in Fig.7 and Fig.8.

During the single satellite test, there was no shielding or interference, so the overall temperature was lower. During the three-satellite test, satellite 2 had the highest temperature because it was in the middle position, and satellite 3 had the lowest temperature because part of the heat dissipation surface on plane +X was opposite to heat sink. The maximum temperature deviation of the three satellites from each other and from the single satellite was not more than 2℃.

3) Analysis of cooling rate in thermal vacuum test

The multi-satellite thermal vacuum test mainly involves three aspects of temperature rising rate, temperature falling rate and temperature stability control[10], among which the temperature rising rate and temperature temperature-property are basically the same as those of a single satellite, which will not be described in detail in this paper. For cooling, due to the influence of occlusion, there is a partial effect.

Analyze the cooling rate of the three- satellite simultaneous test and the single- satellite test, and use a uniform initial temperature of 25℃ to start cooling. Take the battery as an example, as shown in Fig.9.
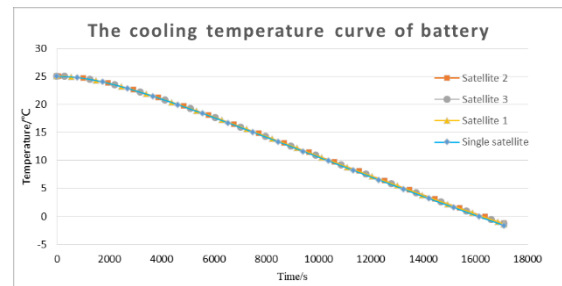
Fig.9. The cooling temperature curve of battery

Fig.9 shows that the cooling process of the three satellites in the simultaneous test is basically the same as that of the single satellite test, and the cooling rate of the single satellite test represented by the blue curve is slightly faster.

From the data, the average cooling rates of a single satellite and the three satellites tested simultaneously from 25°C to 0°C were 5.559°C/h, 5.505°C/h, 5.489°C/h and 5.506°C/h, respectively. The cooling rate of single satellite is the fastest, but it is only 0.070℃/h faster than that of satellite 2 at the same time. Therefore, it can be concluded that the time for a single test conducted with three satellites at the same time will not be significantly longer than that for a single satellite.

The theoretical analysis provided the idea of three-satellite simultaneous test and the principle of reducing the difference between the three-satellite simultaneous test and the single-satellite test, and determined the basic scheme of the test. The results of the simulation analysis further indicated the feasibility of three-satellite simultaneous test and provided the basis for the subsequent test.

III. TEST AND FLIGHT VERIFICATION

A. Ground Environment Test

The multi-satellite simultaneous test was verified in a certain type of positive sample thermal test. The test satellite has a body size of about 1m×1m×2m, and the test was carried out in a KM3 space environment simulator. The satellite layout in the chamber is shown in Fig.4. In the test, infrared heating cage was used to simulate external heat flow, and four high and low temperature cycles were completed[11]. After

each rise and drop to the expected temperature, the test entered the stage of high and low temperature maintenance.

1) Thermal vacuum temperature control

In the thermal vacuum test, three satellites all choose battery temperature as reference. Fig.10 shows the rise and fall curves of the three satellite batteries in a test.
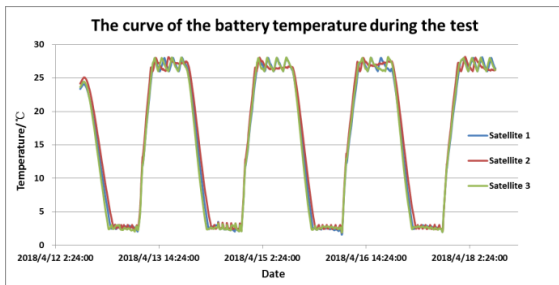
Fig.10 The curve of the battery temperature during thermal test

It show that each thermal test goes through four cycles of high and low temperature, and the temperature curves of the three satellites tested at the same time have a good coincidence, indicating that the rising and cooling rates of the three satellites are basically the same, the temperature control of the three satellites simultaneously meets the requirements of the test specification, and the scheme has passed the test verification.

2) Statistics of satellite test time of each group

5 group satellites (3×5 satellites) had been completed thermal test base on the multi-satellite test method. Because of different test start time, the each group test duration is a little different. The five groups of satellite thermal vacuum tests took 7 days, 7 days, 8 days, 8 days and 8 days respectively, a total of 38 days. If a single six-day test was conducted on each satellite alone, it would take 90 days. It can be seen that the simultaneous test with three satellites could save 57.8% of the test time, and the number of tests could be reduced from 15 to 5.

B. On Flight Verification
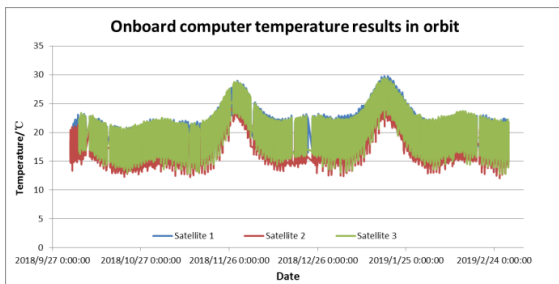
Fig.11 Onboard computer temperature results in orbit

At present, all in-orbit satellites verified by multi-satellite simultaneous test work well, and their performances meet the requirements, among which, the computer temperature is shown in Fig. 11. Moreover, many satellites have reached the design life and extended service stage successively, and the satellites work normally. The on-orbit flight shows that the simultaneous thermal test of three satellites can fully assess the performance of the satellite under vacuum thermal environment.

IV. CONCLUSION

At present, the satellite industry is developing towards miniaturization and mass production[12]. Efficient vacuum thermal test is an effective way to meet the rapid and flexible production and development requirements of batch satellites, reduce the development cycle and increase the efficient cost ratio.

- The analysis shows that the larger the view factor of the satellite to heat sink, the smaller the ratio of the satellite surface area to heat sink surface area, the lower the satellite surface emissivity, the smaller the temperature difference between the satellites under the same heat flow condition, and the smaller the temperature deviation from the single satellite independent test.

- Under the condition of existing satellite surface characteristics, reasonable arrangement and orientation of satellites can effectively reduce the differences between satellites, thus improving the consistency of temperature field and test accuracy of thermal balance test, and reducing the difficulty of temperature control during thermal vacuum test.

- The method of multi-satellite simultaneous test has been verified on a certain project, and the test cost, including funds, labor and duration cost have been greatly reduced, which greatly improves the efficiency ratio of the test.

The multi-satellite test method presented in this paper can be used as a reference for each kinds of small satellite test verification.

REFERENCES

[1] Ma Youli. The launch satellite should do thermal vacuum test [J]. Ring mode technology. 1999,1(58):1~9

[2] Zhang Jingchuan. Application status and development trend of spacecraft vacuum thermal test and control system[J]. Spacecraft Environmental Engineering, 2012,35(3):263~267

[3] Zhou Yingqing, Luo Chao,et al. Development Trend and Thinking of Commercialized Micro Satellite Industry[J]. International space, 2016,05(26):37~43

[4] Hu Xingyi. Increasing the cost-effectiveness ratio of the space segment system and operation of space science missions[J]. Satellite and Internet, 2015,9:68~75

[5] Yang Xiaoning, Gong Jie. Feasibility study on parallel test of satellite antenna thermal vacuum environmental strain test[J]. Spacecraft Environmental Engineering, 2008, 25(4):334~337

[6] Yang Shinming, Tao Wenquan. Heat Transfer [M] (Fourth Edition), Beijing: Higher Education Press, 2006:401~408

[7] HUANG G P. Discussion of the temperature related test technology in thermal vacuum equipment[J]. Environmental Testing, 2012(2): 21-26

[8] Liu Zhonghua, LI Shujie, LIU Guoqiang. Temperature control in thermal vacuum test equipment Formula Research [J]. Electronic Product Reliability and Environmental Test, 2012, 30(4) : 7-11.

[9] YANG Y, LIU B, WANG J, et al. The thermal vacuum test for telecommunication satellite[J]. Spacecraft Environment Engineering, 2013, 30(3): 275-279

[10] LIU F. Thermal control method for thermal vacuum test[J].Spacecraft Environment Engineering, 2003, 20(3): 1-6

[11] GU M. Parameter analysis of loop temperature control for vacuum thermal test[J]. Spacecraft Environment Engineering, 2010, 27(5): 611-615

[12] He Yizhou. Research on Small Satellite Technology and Industrial Development [J]. Telecommunications Network Technology, 2017 (10) : 29-33.

Satellite Control and Data Processing Unit Software Design Based on Multi-core Processor

Junwang He*
Department of Satellite Software Technology Rearsch
Shanghai Engineering Center for Microsatellites
Shanghai, China
hjw10407@163.com

Luyang Xu
School of Software Engineering
University of Science and Technology of China
Hefei, China
xuluyang@mail.ustc.edu.cn

Dongxiao Xu
Department of Satellite Software Technology Rearsch
Shanghai Engineering Center for Microsatellites
Shanghai, China
xudongxiao2016@163.com

Shunjing Yu
Engineering Committee
Shanghai Engineering Center for Microsatellites
Shanghai, China
yusj@microsate.com

Kaige Wang
School of Information Science and Engineering
Harbin Institute of Technology, Weihai
Weihai, China
160200922@stu.hit.edu.cn

Liang Chang
Department of Satellite Software Technology Rearsch
Shanghai Engineering Center for Microsatellites
Shanghai, China
changl@microsate.com

Abstract—**Satellite control and data processing unit is the core of satellite information system, which is the key to realize autonomous flight, autonomous management and improve reliability and security of satellite. Currently, the improvement of the performance of single-core processor is restricted by Moore's law and the power consumption wall. The computing power of the satellite control and data processing unit using a single-core processor can no longer meet the requirements of the satellite's increasing control frequency, autonomous operation capability, and fault diagnosis capability. Based on the research of multi-core processors and the analysis of satellite control and data processing systems, this paper proposes an architecture of satellite control and data processing unit embedded software system based on multi-core processor, which significantly improves the processing performance of satellite information systems. Based on this architecture, it is possible to further improve the autonomous operation capability and fault diagnosis capability of the satellite system.**

Keywords—*multi-core, satellite control, satellite data processing, software architecture*

I. INTRODUCTION

Today, developing software running on satellite control and data processing unit to realize orderly management and integrated data processing of various components of satellites, especially micro satellites, is the key point of satellite software development. With the increasing requirements of satellite system in terms of control frequency, autonomous operation ability, fault diagnosis ability and other aspects, the single-core processor, restricted by Moore's Law and influenced by power wall factors, cannot further improve its computing power when its power is limited, and gradually becomes the bottleneck of satellite control and information system design. In addition, the power consumption of satellite control and data processing unit is particularly prominent compared with the system on the ground due to the resources and environment constraints of satellite system.

Multi-core processors have greater advantages in computing power and energy consumption. Therefore, the use of multi-core processor in the design of satellite control and data processing units is an inevitable choice to improve the processing capacity of satellite information systems.

However, because of high concurrency and resource competition, it is difficult to develop software systems on multi-core processor. There are mainly the following aspects:

1) Parallel programming of multi-core processors is different from single-core processors. The tasks processed by the multi-core processor are executed by multiple cores simultaneously, and the execution sequence need reasonable arrangement and comprehensive design by software developers.

2) Due to the high concurrency of multi-core processors, complete management are required for access to various system resources.

3) Due to the high real-time requirement of satellite control and data processing, the operation timing sequence of each core should be planned as a whole and synchronization is needed in proper moment.

4) It is more difficult to debug parallel programs of multi-core processors than that of single core processors. The time sequence of multi-core processor processing tasks is variable, in which the conditions that may cause program errors may not be fixed each time, that is, the recurrence rate of the same location is low, which increases the time for developers to debug.

Therefore, we need to fully consider the inter-core mechanism in the software system design to realize perfect and orderly data and peripheral management [1]. On this basis, multi-core task allocation and operation timing design are reasonably carried out, so that the high real-time and timing dependency of satellite control and management can be satisfied.

II. MULTICORE MANAGEMENT MECHANISM

A. Multi-Core System Architecture

The Multi-core System Architecture [2] is shown in Fig. 1. Each core runs an operating system separately, and each operating system manages different applications. Data sharing and synchronization between cores can be managed by a multi-core unit. If a large amount of data is to be transmitted, efficiency can be improved by sharing memory.

---

*Junwang He is the corresponding author. (e-mail: hjw10407@163.com).
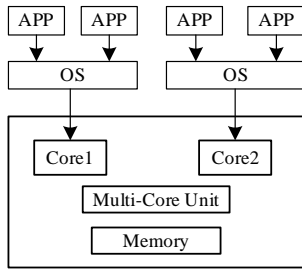
Fig. 1.   Multi-core System Architecture

B. *Multi-core Unit Design*

Multi-core can realize parallel execution of programs. By distributing tasks to different cores, the execution efficiency of system tasks can be greatly improved, thus increasing the control frequency of the system. Although multi-core can greatly improve the execution efficiency of the system, it also increases the complexity of the system. The problem of communication and synchronization between cores should be solved first. Therefore, it is necessary to design a Multi-Core Unit for inter-core data transmission and inter-core control [2]. The structure of the multi-core unit is shown in Fig. 2.
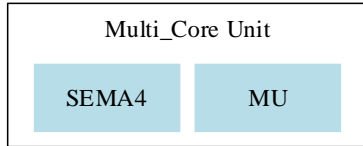
Fig. 2.   Multi-Core Unit

C. *Message Unit*

The messaging Unit (MU) module enables two cores in an SoC to communicate and coordinate messages (for example, data, state, and control) via the MU interface. The MU also provides the ability for one core to signal the other core using interrupts [2].

The specific design of MU module is shown in Fig. 3. The MU module is divided into two sides. Each side is connected to the processor through the bus and interrupt controller. Each register is 32-bit wide. TX/RX registers are used to transmit and receive data information between cores. The state and control register is used to show the transmitting and receiving state of data, and also to control the masked state of MU universal interrupt [3]. The interrupt module can generate interrupt requests to notify the core that data is available to transmit or receive.
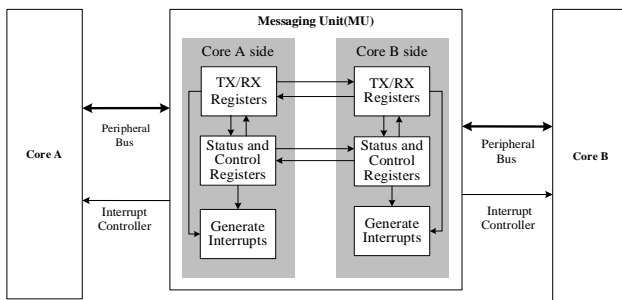
Fig. 3.   Message Unit Model

D. *Semaphores*

Multi-core processors can realize parallel execution of programs, thus greatly improving the execution efficiency of programs. However, parallel execution of programs will cause race conditions, so we need a secure locking mechanism to control access to shared data structures, shared hardware resources, and shared memory. Software can use these locking mechanisms to achieve synchronous access to shared data or resources, prevent race conditions, and maintain memory consistency.

In a single core processor system, a variable can be defined to implement the locking mechanism. When different threads need to access the shared resources, they can first access the variable to decide whether to obtain the right to access the shared resources. However, unlike the lock mechanism within a single core, the lock mechanism between cores is difficult to implement with simple variables, so a hardware implemented lock is very necessary.

The Semaphores provides robust hardware support needed in multi-core systems for implementing semaphores and provides a simple mechanism to achieve "lock and unlock" operations via a single write access [3]. The Semaphores module implements hardware-enforced semaphores as an IPS-mapped slave peripheral device as shown in Fig. 4.
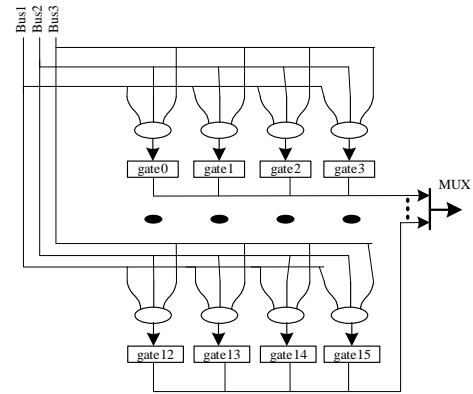
Fig. 4.   Semaphores block diagram

The feature set includes[3]:

1) Once locked, the gate can and must be unlocked by locking core.

2) The number of implemented gates is specified by a hardware configuration define.

3) Cores lock gates by writing "core_number+1" to the appropriate gate and must read back the gate value to verify the lock operation was successful.

4) Each hardware gate appears as a 16-state, 4-bit state machine. If the gate value is 0, the state is unlocked, and if the gate value is n(n>0), the lock is used by the core (n-1).

E. *Shared Memory*

In a multicore system, different tasks can be assigned to different processor cores in order to improve the operating efficiency of the system. For example, more time-consuming programs such as peripheral data acquisition can be executed

on cores with weak computing power, while processor cores with strong computing power are dedicated to processing these data and executing some algorithm which demands large computing power. Pipeline operation can greatly improve the operating frequency of the system. Message unit can be used to share a small amount of data between cores. When the amount of data exceeds 10K bytes, the data sharing efficiency of message unit will be greatly reduced. Therefore, another more efficient data sharing method is necessary: shared memory.

Shared memory is the fastest of all communication mechanisms, because it can read and write shared memory directly without sending byte by byte, which saves a lot of time, and the size of shared memory can be customized. A synchronization mechanism is required for access to shared memory between processes in a single core system, as well as for shared memory between multiple cores. Compared with other methods, interrupt is more suitable for multi-core synchronization, as shown in Fig. 5.
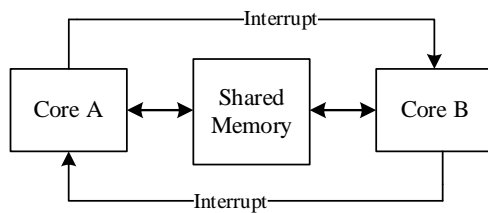
Fig. 5.   Shared memory model

In order to further improve the data transfer efficiency of shared memory between multi-core, we adopt the mechanism of no-copy-send and no-copy-receive, as shown in Fig. 6. The sending and receiving steps are as follows [3]:
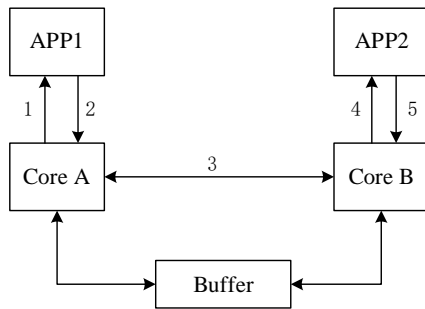
Fig. 6.   Non-copy send and receive

1) The application on core A obtains the shared buffer pointer allocated by the driver and fills in the data to be sent in this memory.

2) Notify the kernel that the data is completed and ready to be sent.

3) Processor A notifies processor B to prepare to receive data through interruption and sends the address pointer of the buffer to B.

4) Application on core B obtains the address pointer of the buffer and directly reads the data of the buffer.

5) Processor B releases the buffer in preparation for the next data transfer.

III. TASK DISTRIBUTION AND TIME SEQUENCE DESIGN

A. Task Overview

Satellite control and data processing unit is composed of multiple subsystems, including satellite service management, attitude control, telemetry and remote control, energy management, thermal control, status acquisition, et al. Each subsystem function is implemented as a thread, and the operation timing is related to the importance and dependency of the subsystem, the important thread would run first. In order to ensure the order and effectiveness of the arrangement of each subsystem, satellite service management use semaphores to strictly controls the timing of each thread. The first thread which runs in satellite control and data processing unit is the satellite management thread, whose main function is to ensure that the other sub-system threads running normally, followed by the data acquisition thread, which collects all kinds of sensors carried by the satellite to sense the satellite state and provide inputs for threads follow up. There are RS422 asynchronous serial port, synchronous serial port, can, RS485 asynchronous serial port, et al. on the communication bus between the sensor and the satellite. The communication architecture is request-response mode. The advantage of this method is that passive transmission does not always occupy the communication bandwidth and is more efficient. The shortcoming is also obvious. There is a time interval between the request and the response. It takes time for the satellite to send the telemetry request, and the sensor response takes more than 10ms. For example, the 422 serial port can wait in parallel, saving time; but CAN bus nodes waits in serial waiting, when the sensor on the satellite communicates in CAN bus mode, the waiting time will be very considerable, and the dual-core system can solve the above problems.

B. Task Distribution

Dual-core and multi-core processors, especially processors that can run asynchronously with multiple cores, can achieve true multi-threading, rather than multi-threading implemented by the operating system through thread suspend and resume switching. The advantages of the multi-core system are self-evident. The disadvantage is that when the task is executed in the multi-core system, the communication delay caused by the inter-core communication is 3 to 5 times that of the intra-core communication [4-5]. In this paper, dual-core system is studied, which can make full use of its advantages and avoid its disadvantages by rationally allocating dual-core tasks according to the characteristics of satellite operations and their respective priorities.

The dual-core processor selected in this paper includes a Cortex-M4 core and a Cortex-M0+ core (respectively Core A and Core B). Cortex-M4 has an FPU, which handles floating-point operations more than 100 times that of Cortex-M0+, suitable for Floating-point operations, such as the four-element solution of attitude control threads; Cortex-M0+ does not have FPU, so it is suitable for integer operations and time sequence control. Both cores can perform data interaction by enabling peripherals, and have the right to enable the same peripheral.

Based on the characteristics of the satellite tasks and dual-core, the collection task is suitable for Core B, collect sensor data through the bus, and send instructions to the sensor; other tasks than the collection task run in Core A,

Core A manages Core B through inter-core communication [6], Core B sends the collected data to Core A. Core A performs the follow up tasks on the satellite. Core A and Core B run asynchronously. Core A uses the data collected by Core B in the previous cycle and enables the cycle Core B collection. A and Core B are running simultaneously to achieve high efficiency.

C. Time Sequence and Synchronization Design

In order to run dual-core efficiently and reduce the waiting time of one Core A and Core B, dual-core operation timing sequence should be rationally designed. Take Fig. 7 as an example to introduce the typical operation timing of Core A and Core B.

Among them, C1: Core A requests Core B data, C2: Core A enables Core B to collect data, C3: Core A completes data processing in this cycle and transmits data under telemetry. R1: Core B return data response, R2: Core B peripheral telemetry receive complete response. The meaning of each time period in the Fig. 7 is shown in TABLE I.

T0, T1, T2 and T3 should have the relationship which is showed in Equation 1.

$$T0 \geq \max(T1, T2 + T3), T0 = \frac{1000}{n}, n \in \{1,2,3, \dots\} \quad (1)$$

The unit of time T0 of each cycle is millisecond. In order to ensure the entire second operation of the satellite, formula 1 is adopted.

- Core B running process

Core B responds to three interrupt requests, C2 interrupt request, C1 interrupt request, C3 interrupt request, and the priority of them is decreasing in order. Define two types of shared memory: Core A and Core B. Shared memory 1: Core B sent sensor data to Core A; Shared memory 2: Core B sent remote control data to Core A. When Core B receives the Core A C2 request, it collects the sensor data in the main loop and writes the data to Shared memory 1. When Core B receives an interrupt request from Core A C1, Shared memory 1 is sent. When Core B receives the earth station remote control instruction, it fills the Shared memory 2 with the instruction data and sends the Shared memory 2 to Core A.
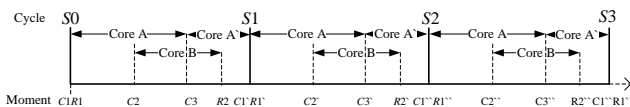
Fig. 7.   Typical timing for dual-core operation

TABLE I.   TIME PERIOD MEANING

Period	Meaning
C1~C2	Core A delayed acquisition time
R2~R1`	Core B reserved time
C1~C3	Core A data processing time T1
C2~R2	Core B runtime T2
C1~C1`	Cycle time T0
C3~C1`	Core A telemetry download time T3

- Core A running process

All threads except the data acquisition/collection thread are running on Core A, and the operating system is responsible for managing each thread in the same way as the satellite control and data processing software used in single-core processor. Among them, the operation time T1 of Core A and the telemetry downlink time T3 of Core B have a sequential logical relationship and cannot overlap. Among them, the operation time T1 of core A and the downlink time T3 of telemetry data of Core B have a sequential logical relationship and cannot overlap. The request C2 sent by Core A to Core B should be before C1'. Moving C2 to the right can improve the timeliness of core B collecting sensor data. Analyzing Fig. 7 and Equation 1, we can calculate the optimal dual-core operation under the following three different conditions. According to the principle of highest efficiency, we can reasonably arrange the dual-core timing.

The situation shown in Fig. 8 depends on Core B running time. Core B running time T2 is greater than Core A running time T1 and Core B telemetry down pass time T3.
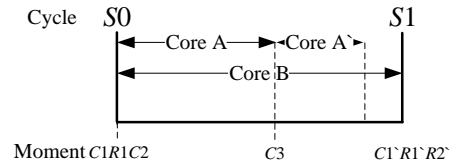
Fig. 8.   Dual-core operation timing 1
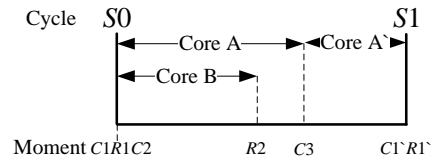
Fig. 9.   Dual-core operation timing 2
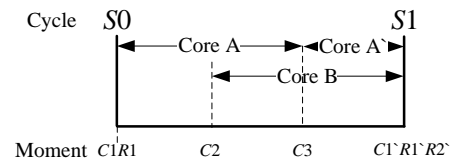
Fig. 10.   Dual-core operation timing 3

The situation shown in Fig. 9 depends on Core A operation time T1 and Core B telemetry data transmission time T3, and the moment when Core A notifies Core B to acquire data is moved to the beat start moment. In this way, it is not necessary to judge the starting timing of Core B, and the situation that the response time R2 for receiving completion from the core exceeds the next beat can be avoided.

Fig. 10 shows the ideal situation of dual-core task arrangement. The master core starts the acquisition of data from the slave core at an appropriate time, so that the acquisition completion time of the slave core coincides with the beat end time, thus realizing the most efficient data acquisition and the highest dual-core operation efficiency.

In summary, the starting timing of each moment should be reasonably arranged based on the principle of highest efficiency, combined with the best timeliness and considering the time margin, as shown in Fig. 7.

IV. IMPROVED PERFORMANCE

Table II shows measured data of processing margin from a satellite under development which use AT697F processor in control and data processing unit. AT697F is an radiation-hard single-core processor which is widely used in a lot of spacecrafts.

TABLE II. A MEASURED TIME ON SINGLE-CORE PROCESSOR

Cycle ID	Typical time used	Margin
A	109ms	16ms
B	77ms	48ms
C	87ms	38ms
D	88ms	37ms
E	79ms	46ms
F	79ms	46ms
G	77ms	48ms
H	75ms	50ms

TABLE III. AN EVALUATED TIME ON DUAL-CORE PROCESSOR

Cycle ID	Typical time used	Margin
A	59ms	66ms
B	47ms	78ms
C	57ms	68ms
D	58ms	67ms
E	49ms	76ms
F	49ms	76ms
G	57ms	68ms
H	55ms	70ms

If we use the designed software architecture on a dual-core processor whose core has similar performance with the single-core of AT697F, we could achieve a evaluated processing margin as Table III. Nearly, a control frequency of 16Hz could be achieved while 8Hz control frequency on single-core processor. Then, the attitude stability performance of satellite could be improved due to a higher control frequency.

V. CONCLUSION

This paper designs an architecture of satellite control and data processing unit embedded software system based on dual-core system. Compared with the single-core processor system, it distribute the tasks based on characteristics of core and requirement of each task, giving full play to the advantages of the dual-core processor, improving the performance of the satellite control and data processing software, and having certain reference significance for the design of other multi-core satellite computer system.

REFERENCES

[1] Kriaa, Lobna,Bouchhima, Aimen,Gligor, Marius,Fouillart, Anne-Marie,Pétrot, Fréderic,Jerraya and Ahmed-Amine, "Parallel Programming of Multi-processor SoC: A HW-SW Interface Perspective," International Journal of Parallel Programming, 2008, 36(1).

[2] NXP, RV32M1 Series Reference Manual, Document Number: RV32M1RM, Rev. 1.1, November 2018

[3] NXP, Develop a Multicore Application for RV32M1-VEGA, Application Note, Rev. 0, October 2018.

[4] Wang Y , Shao Z , Chan H C B , et al. Memory-Aware Task Scheduling with Communication Overhead Minimization for Streaming Applications on Bus-Based Multiprocessor System-on-Chips[J]. IEEE Transactions on Parallel & Distributed Systems, pp. 1797-1807, July 2014.

[5] Xiaojie Xu,Center N,Lisheng Wang.Task assignments based on shared memory multi-core communication[C] //Proceedings of the International Conference on Systems and Informatics.Shanghai,China: IEEE,2014.

[6] WU Li-sheng.Research and Implementation of Embedded Multi-core Operating System Kernel Based on ARM Cortex-A9 MPCORE. Digital Technology &Application, pp. 134-135, March 2020.

Analysis of Ionospheric Scintillation Detection Based on Machine Learning

Lin Mengying
School of Instrument Science and Engineering
Southeast University
Nanjing, China
230198865@seu.edu.cn

Zhu Xuefen*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
220193298@seu.edu.cn

Luo Yimei
School of Instrument Science and Engineering
Southeast University
Nanjing, China
zhuxuefen@seu.edu.cn

Yang Fan
School of Instrument Science and Engineering
Southeast University
Nanjing, China
582919925@qq.com

Abstract—**Ionospheric scintillation, characterized by deep fluctuation of signal intensity on amplitude and phase, poses threats to the global navigation satellite system (GNSS) tracking performance, even causing degradation in position and navigation in accuracy, integrity and continuity. To achieve higher performance of ionospheric scintillation detection, a method of machine learning is implied to attain automatic detection. Three levels of scintillation intensities and two types of scintillation features are investigated and compared. Based on data from São José dos Campos, Brazil (23.2S, 45.9W), training validation and prediction results show 99% accuracy approximately and above 96%, respectively. The analysis will provide useful guidance on ionospheric scintillation detection to make of great signification on design of advanced GNSS receiver and research of atmospheric layer and space weather.**

Keywords—ionospheric scintillation, machine learning, scintillation intensity, scintillation feature

I. INTRODUCTION

Ionospheric scintillation, refers to the phenomenon of rapid fluctuation of signal intensity and phase jittering when the radio frequency signals pass through the ionospheric electron density irregularities[1][2]. Normally, it frequently occurs in equatorial, auroral, and polar regions where particularly suffused with free electrons with long duration. Numerous factors will conduct in ionospheric scintillation, including solar activity, geomagnetic storm, local electric field, electrical conductivity, wave interaction and so on[3].

In the equatorial region, occurrence of strong ionospheric scintillation will pose a threat to the global navigation satellite system (GNSS) receiver tracking performance, including cycle slips, phase errors and increased carrier Doppler shifts [4]. In severe cases, the receiver measurements which the deep simultaneous signal fading during ionospheric scintillation on all three signal bands rarely occurs, will be plagued with serious destruction, causing in deep amplitude fading and a series of errors on carrier tracking loop, even leading to a degradation in position and navigation solution accuracy, integrity and continuity [5]. The loss of lock on carrier tracking loop, caused by increased Doppler frequency shift during equatorial ionospheric scintillation, may occur when the total Doppler frequency shift exceeds the preset bandwidth of tracking [6].

As two types of representative scintillation indicators used to describe the intensity of ionospheric scintillation occurrence, amplitude scintillation implies the signal-to-noise ratio (SNR) drop rapidly and the signal interruption. Its corresponding amplitude scintillation index S_4 is described as the standard deviation of the normalized signal intensity. Typically, there will be regarded as weak ionospheric scintillation occurrence during the signal propagating through the ionosphere, when S_4 index is below 0.2. Similarly, the value of S_4 index between 0.2 and 0.5 will be considered as medium ionospheric scintillation. Once S_4 is above 0.5, we consider as strong scintillation event which frequently occurs around the magnetic equatorial region between local sunset and midnight and especially most intense during the period of solar maximum year [7][8]. Another typical index σ_ϕ is defined as the standard deviation of the detrended carrier phase measurements, normally leading to phase cycle slip and even the loss of lock of signal to reduce the receiver performance in carrier tracking loop [9][10].

A series of efforts and researches have made to decrease the interferences on the systems that rely on radio frequency signals passing through the atmospheric layer, such as GNSS. Particularly, the problems of how to detect and eliminate the factor of ionospheric scintillation, have been drawn much attention in scientific field as well as industry. In the previous researches, the typical two scintillation indicators S_4 and σ_ϕ are normally choose to identify the occurrence amplitude and phase scintillation. Once the observation exceeds the predicted value of S_4 and σ_ϕ, there will be regarded as scintillation event occurring. Moreover, selections of various detrending methods as well as corresponding parameters on process of scintillation calculation will affect the final outputs of S_4 and σ_ϕ, such as discrete wavelet filtering, continuous wavelet filtering, Butterworth filtering and polynomial fitting [11][12]. Based on above analysis, apart from two main scintillation indictors, it will be more convincing and suitable to take more scintillation features into consideration on the detection of scintillation events.

According to above analysis on previous detection methods, the support vector machine (SVM) model is proposed to detect the events of amplitude ionospheric scintillation, with GPS L1 data collected in equatorial region

located at São José dos Campos, Brazil (23.2S, 45.9W) from 2013 to 2015. Moreover, with more scintillation features being selected to train the model to overcome the drawbacks, SVM models based on two types of features will be investigated.

The rest of the paper is organized as follows. Section 2 introduces the data collecting process and a brief description on selection of ionospheric scintillation samples, as well as the detection model. Section 3 presents SVM training results based on samples of three levels of scintillation and two types of features, with corresponding comparisons analyzed. A conclusion and recommendations for future work are given in section IV.

II. METHODOLOGY

A. Support Vector Machine

As one of the most widespread used model in machine learning algorithms, SVM has been utilized in various research fields as a technique of classification and regression. The core of SVM is to find the optimal hyperplane by given samples with some features to achieve classification, further being applied to predict and classify novel data. There are several types of kernel function can be configured in SVM model, such as linear kernel, back propagated neural network (BPNN) kernel and radio basic function (RBF) kernel which can be also called Gaussian kernel [8]. Four features of ionospheric scintillation will be considered and the RBF kernel $K\langle x, y \rangle = \exp\left(-\dfrac{(x-y)^2}{\gamma}\right)$ is selected in the SVM model here, where γ represents the hyper-parameter of kernel scale.

Assuming that m groups of training data set $\{\mathbf{x}^{(i)}, y^{(i)} \mid i = 1,...,m\}$ are given, while $\mathbf{x}^{(i)} \in \mathbb{R}^n$ represents ith features set with four elements which means that $n = 4$. $y^{(i)} = \{1, -1\}$ stands for the label of scintillation event with 1 for scintillation occurrence and -1 for non-scintillation. The m labels should be labelled manually and then used to train the SVM model for the first time. After optimum SVM classification being trained, the remaining data set without training can be utilized to predict corresponding labels which reflect whether ionospheric scintillation events occur. The SVM model can be established as $y = \mathbf{w}^T\mathbf{x} + b$, where $\mathbf{w}$ and b are variables to decide the optimum detection boundary. The process of finding the optimum hyperplane can be transformed into maximum the decision boundary and described as following mathematical solution.

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C\sum_{i=1}^{m} \xi_i \tag{1}$$

$$\text{subject to} \begin{cases} y^{(i)}\left(\mathbf{w}^T\mathbf{x}^{(i)} + b\right) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases}, i = 1,...,m$$

where $\mathbf{w} \in \mathbb{R}^2$, $b \in \mathbb{R}$ and ξ_i represent the weights, bias and slack variables of each training data set, respectively. C stands for the hyper-parameter which indicates the largest tolerance on training data points exceeding the boundary to prevent over

fitting. After combining with the Lagrange multiplier and dual form, the problem in (1) can be transformed as follows.

$$\min_{\alpha_i} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y^{(i)} y^{(j)} \left(\mathbf{x}^{(i)}\right)^T \mathbf{x}^{(j)} - \sum_{i=1}^{m} \alpha_i \tag{2}$$

$$\text{subject to} \begin{cases} \sum_{i=1}^{m} \alpha_i y^{(i)} = 0 \\ 0 \leq \alpha_i \leq C \end{cases}$$

$$\text{hidden condition} \begin{cases} \mathbf{w} = \sum_{i=1}^{m} \alpha_i y^{(i)} \mathbf{x}^{(i)} \\ \beta_i = C - \alpha_i, i = 1,...,m \end{cases}$$

where the coefficients $\alpha_i, i = 1,...,m$ can be solved by QP function in Matlab. Thus, the satisfied value of $\mathbf{w}$ can be obtained according to the hidden condition. Meanwhile, coefficient b can be also gotten by substituting the data point that is support vector in the detection boundary into the formula of SVM model.

B. Data Collection System

All of data set used here are based on the data collected during the last peak of the solar cycle at the low latitude site of São José dos Campos, Brazil (23.2S, 45.9W) from 2013 to 2015. The collection system location is shown in Fig. 1. Since 2012, a GNSS data collection system has been collecting segments of raw GNSS (multiple frequency/ constellation) IF data at periods of high ionospheric scintillation. The GTEC free front-end are exploited to record the GPS IF signals by generating zero intermediate-frequency data streams with 8 bits resolution I/Q samples at 20MHz complex sampling rate [13]. In this paper, 50 segments of raw IF GPS L1 data in three years from 2013 to 2015, are processed to train SVM models and predict new scintillation events.

In equatorial region, the amplitude scintillation index S_4 is considered as important element to capture the scintillation characteristics. Besides, the carrier frequency will be also regarded as another factor, which the distribution disturbance is affected by different intensities of ionospheric scintillation. Fig. 2 shows an example of distribution between the standard carrier frequency and S_4 with PRN 1, 11, 31 and 32 at 01:00:19-02:00:19 on January 31st, 2013. It indicates that the frequency measurement performs approximately linear relation with corresponding S_4 overall. The S_4 can be obtained as follows.

$$WBP = \sum_{i=1}^{M}\left[\left(I + N_i\right)^2 + \left(Q + N_i\right)^2\right] \tag{3}$$

$$NBP = \left[\sum_{i=1}^{M}\left(I + N_i\right)\right]^2 + \left[\sum_{i=1}^{M}\left(Q + N_i\right)\right]^2 \tag{4}$$

where M represents the number of noise N_i as well as I / Q samples used to measure one wide band power WBP and narrow band power NBP. Furthermore, the difference of NBP and WBP is calculated to explore the connection between NBP, WBP and the raw signal intensity SI_{raw} [11].

$$SI_{raw} \propto NBP - WBP$$
$$= \left(M^2 - M\right)\left(I^2 + Q^2\right) + 2\left(M-1\right)*$$
$$\left(I+Q\right)\sum_{i=1}^{M} N_i + 2\sum_{i \neq j} N_i * N_j \qquad (5)$$
$$\approx \left(M^2 - M\right)\left(I^2 + Q^2\right)$$

The value of M is set to 40 for that GPS navigation data rate is 50Hz, and normalized standard deviation S_4 is defined as the amplitude scintillation index SI_{raw} shown as following.

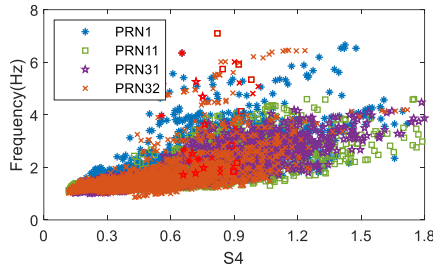Fig. 1.   Location of the data collection system on the map

Fig. 2.   The distribution between the standard carrier frequency and S4 with PRN 1, 11, 31 and 32 at 01:00:19-02:00:19 on 1/31/2015.

$$S_4 = \sqrt{\frac{\langle SI_{norm}^{2} \rangle - \langle SI_{norm} \rangle^2}{\langle SI_{norm} \rangle^2}} \qquad (6)$$

where $\langle \cdot \rangle$ stands for the average value over the interval. The interval time is set to 10s and combines with the moving window for 1s.

Based on 1Hz of S_4 output frequency, all the data are divided into 1mintue blocks without overlap, and corresponding mean value $S_{4,mean}$ and maximum value $S_{4,max}$ of are regarded as two features of scintillation in each block. Similarly, the corresponding carrier frequencies after standardization can be also processed based on same block partition to output another two features mean value of standard carrier frequency $Freq_{mean}$ and maximum value $Freq_{max}$. Thus, the features of samples can be described as $\mathbf{x}^{(i)} = (S_{4,mean}^{(i)}, S_{4,max}^{(i)}, Freq_{mean}^{(i)}, Freq_{max}^{(i)})$. It should be mentioned that these features are outputted with the detrending process before standardization to reduce other interference [11][13]. To make comparison, another type of features group without detrending will be trained and test. In this way, the influence of selection of detrending methods and parameters on the outcome values of features can be eliminated.

III. TRAINING AND VALIDATION

A. Sample Selection

The data training and testing in this paper are collected during the last peak of the solar cycle at the low latitude site of São José dos Campos, Brazil (23.2S, 45.9W) from 2013 to 2015. There are all 50 segments of data are available and each of them contains one hour of data with several numbers of PRNs can be observed. To reduce the effect of multipath, the PRNs with elevation above 30° can ben remained and selected. According to the times of $S_4 > 0.2$ exceeding 30 within 1min, the label of scintillation event will be set as 1, and vice versa. In addition, three groups of PRNs are organized based on intensities of scintillation. In each of group, PRNs from each segment of data are selected and several PRNs in certain segment of data are the same to other groups, because there are only one level of scintillation occurring during the data collecting period. TABLE I shows the PRNs for training and testing on three levels of scintillation intensities. On the whole, there are 2832 data points in each level of scintillation group and the radio of scintillation events over non-scintillation in the group of strong scintillation data, is 2.02:1, 1:1.18 for medium scintillation data and 1:3.07 for weak scintillation data approximately.

TABLE I.   PRNs FOR TRAINING AND TESTING ON THREE LEVELS OF SCINTILLATION INTENSITIES

Data	Start Time	PRN		
		Strong	Medium	Weak
3/20/2013	1:00:11	1	11	32
3/25/2013	1:00:01	11	23	20
3/25/2013	2:00:01	1	20	32
3/26/2013	0:00:02	1	31	32
3/26/2013	1:00:02	11	20	32
3/26/2013	2:00:03	1	1	23
7/18/2013	21:00:14	8	7	23
10/13/2013	23:59:59	24	2	12
10/22/2013	23:00:08	2	24	12
10/23/2013	1:00:07	29	12	12
10/24/2013	23:00:07	2	24	12
10/27/2013	22:59:59	24	2	12
10/27/2013	23:59:59	29	5	12
10/29/2013	1:00:02	25	12	29
10/30/2013	2:00:04	25	21	29
11/18/2013	0:00:01	25	29	29
11/18/2013	1:00:01	25	21	21
11/18/2013	23:00:02	5	29	25
11/23/2013	0:00:08	21	29	29
11/28/2013	0:00:09	18	21	29
11/28/2013	1:00:08	18	29	21
12/4/2013	23:00:05	21	25	29
12/9/2013	23:00:02	25	29	21
12/11/2013	23:00:05	25	21	29
12/12/2013	0:00:05	18	29	21
12/12/2013	23:00:11	25	29	29
12/13/2013	23:00:08	25	29	21
12/13/2013	0:00:10	18	21	21
12/14/2013	0:00:07	29	18	21
12/15/2013	23:00:01	18	25	21
12/16/2013	0:00:01	18	29	21
12/18/2013	23:00:05	18	21	29
1/30/2014	1:00:04	14	22	22
1/30/2014	0:00:04	14	22	22
2/25/2014	23:59:41	11	19	32
2/27/2014	1:00:03	1	32	32
2/27/2014	0:00:04	19	31	31
3/3/2014	23:00:07	3	19	14
3/3/2014	0:00:06	31	11	32
3/9/2014	23:59:59	1	31	31
3/12/2014	0:00:01	11	31	31
3/23/2014	23:00:10	11	32	32
3/24/2014	0:00:10	1	31	23
11/19/2014	0:00:09	25	21	21
11/27/2014	0:00:11	18	29	21
12/20/2014	0:00:12	22	18	21
1/31/2015	1:00:19	19	32	4
2/7/2015	0:00:09	32	4	14

Before SVM model training, k folds validation is configured which means that In the training process, the input training samples will be divided randomly into k parts with each k-1 parts for training model and the remainder for testing

the trained model to obtain the test accuracy. The value of k will be configured as 10 which is the default.

B. Training and Validation

To evaluate the performance of each SVM classification, the training samples with strong scintillation shown in the 3th column on TABLE I are selected, while the other two samples with medium and weak scintillation(in the 4th and 5th columns) are considered as testing the trained model, respectively and vice versa. On the process of SVM training and validation, the Box Constraint C and Kernel Scale γ are considered as follows: $C=2^{c_0}$, $C_0 \in \{-5, -4.5, -4, ..., 9.5, 10\}$ and $\gamma=2^{\gamma_0}$, $\gamma_0 \in \{-5, -4.5, -4, ..., 4.5, 5\}$. Based on above hyper-parameter ranges, each couple of hyper-parameters for each training samples will be tried to reach the optimization with 10 folds cross validation. After 10 times of training, the current hyper-parameter C and γ as well as corresponding average accuracy will be outputted. Moreover, samples with two types of features will be trained and compared. TABLE II shows the two optimal hyper-parameters C and γ, as well as corresponding validation accuracy. As can be seen, all of the accuracies are above 99% with excellent performances on the condition of three levels of scintillation intensities and two types of features. It can be implied that the training samples of features without detrending process also contain the characteristics of ionospheric scintillation so the trained models can be of sufficient performance.
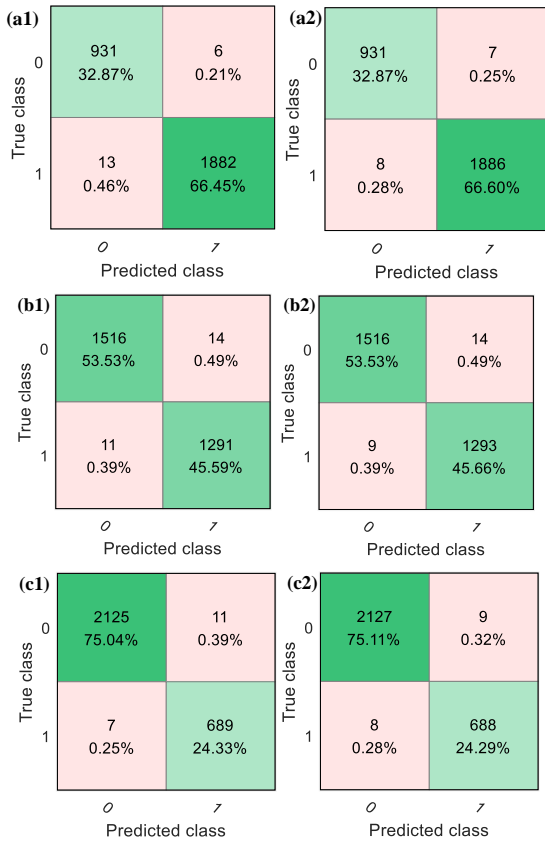

Fig. 3. Performances of training results on RBF kernel SVM models based on three levels of scintillation intensities and two types of features. (a1) Strong scintillation features with detrended process. (a2) Strong scintillation features without detrended process. (b1) Medium scintillation features with detrended process. (b2) Medium scintillation features without detrended process. (c1) Weak scintillation features with detrended process. (c2) Weak scintillation features without detrended process.

Training Results (RBF Kernel)	Scintillation Intensity		
	Strong	Medium	Weak
Features With Detrended Process			
Accuracy	99.33%	99.12%	99.36%
Box Constraint C	256	724	181
Kernel Scale γ	2.8284	8	2
Features Without Detrended Process			
Accuracy	99.47%	99.19%	99.40%
Box Constraint C	91	724	512
Kernel Scale γ	1	4	8

Furthermore, to analyze the details of samples, corresponding confusion matrixes are shown in Fig. 3. On each panel of confusion matrix, the samples number and rate have been marked out . It should be noted that there are also several scintillation events with predict classes marked as 1 can be detected even though the true classes are marked as 0, which means that the detection method proposed in this paper can classify some scintillation events with not above 0.2. This phenomenon will be of great significance on the research of atmosphere and space weather and improvement on design of scintillation detection receiver.

C. Testing and Prediction

On the condition of trained SVM classification models, a series of tests are operated based on two levels of scintillation datasets and two types of without corresponding training datasets. TABLE III shows the prediction performances of SVM models on different training models and testing samples. Comparing two prediction results, on the condition of medium scintillation training samples with detrending process features, the testing samples accuracy with strong scintillation (99.29%) shows slightly less than that with weak scintillation (99.40%), with same phenomenon as condition of features without detrended process. It can be illustrated that the SVM model trained by medium scintillation samples lack enough distinct characteristics to classify some scintillation events from medium and weak scintillation events. Based on trained models with other scintillation intensities as well another type of features, this phenomenon can be also identified.

Especially, as can be seen on the whole, all of the accuracies are above 96.5% with great performance. Moreover, the prediction accuracies achieve 99% approximately ranging from 98.59% to 99.40% based on features with detrending process, while accuracies range from 96.75% to 99.29% based on features without detrending process. It means whether features are detrended, the scintillation events can be classified and detected successfully with high accuracy above 96% based on these processes of training and testing.

Prediction Accuracy		Training Samples		
		Strong	Medium	Weak
		Features With Detrended Process		
	Strong	N/A	99.29%	99.15%
	Medium	98.59%	N/A	99.12%
Testing Samples	Weak	99.26%	99.40%	N/A
		Features Without Detrended Process		
	Strong	N/A	98.45%	99.01%
	Medium	96.75%	N/A	99.01%
	Weak	98.06%	99.29%	N/A

IV. Conclusion

The RBF kernel SVM classification model configured as 10 folds and cross validation is proposed to achieve sample training and testing, according to strong, medium and weak scintillation samples as well as features with/without detrending process. Comparisons on different scintillation intensities and features present great validation and prediction accuracies overall, which are approximately 99% and above 96%, respectively. In the future research on ionospheric scintillation detection, relevant hyper-parameters will be further optimized and more automatic methods will be explored. Meanwhile, more scintillation data of Beidou navigation system as well as data collected from other locations will be utilized to fulfill the sample dataset, which provide helpful guidance on design of scintillation detection GNSS receiver as well as research on atmosphere and space weather.

Acknowledgment

References

[1] K. C. Yeh and C. H. Liu, "Radio wave scintillations in the ionosphere," Proc. IEEE, vol. 70, no. 4, pp. 324-360, April 1982.

[2] J. Aarons, "Global morphology of ionospheric scintillations," Proc. IEEE, vol. 70, no. 4, pp. 360-378, April 1982.

[3] S. Basu, E. MacKenzie and S. Basu, "Ionospheric constraints on VHF/UHF communications links during solar maximum and minimum periods," Radio Science, vol. 23, no. 03, pp. 363-378, May-June 1988.

[4] C. Cristodaro, F. Dovis, N. Linty and R. Romero, "Design of a configurable monitoring station for scintillations by means of a gnss software radio receiver," IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 3, pp. 325-329, March 2018.

[5] Y. Jiao, D. Xu, Y. Morton and C. Rino, "Equatorial scintillation amplitude fading characteristics across the GPS frequency bands," Navigation, vol. 63, no. 3, pp. 267-281, Sep. 2016.

[6] K. Guo, M. Aquino, S. V. Veettil, "Ionospheric scintillation intensity fading characteristics and GPS receiver tracking performance at low latitudes," GPS Solutions, vol. 23, no. 43, Feb. 2019.

[7] E. J. Fremouw, R. L. Leadabrand, R. C. Livingston, M. D. Cousins, C. L. Rino, B. C. Fair and R. A. Long, "Early results from the DNA wideband satellite experiment–complex-signal scintillation," Radio Science, vol. 13, no. 1, Jan. 1978.

[8] Y. Jiao, J. J. Hall and Y. T. Morton, "Automatic equatorial GPS amplitude scintillation detection using a machine learning algorithm," in IEEE Transactions on Aerospace and Electronic Systems, vol. 53, no. 1, pp. 405-418, Feb. 2017.

[9] K. C. Yeh and C. Liu, "Radio wave scintillations in the ionosphere," Proc. IEEE, vol. 70, no. 4, pp. 324-360, May 1982.

[10] D. Y. Xu, Y. Morton, D. Akos, D. Akos and T. Walter, "GPS multi-frequency carrier phase characterization during strong equatorial ionospheric scintillation," Proc. ION GNSS, 2015, pp. 3787-3796.

[11] F. Niu, "Performances of GPS signal observables detrending methods for ionosphere scintillation studies," M.S. thesis, Dept. Computational Sci. and Eng., Miami University, Oxford, USA, 2012.

[12] Y. Liu, Y. J. Morton and Y. Jiao, "Application of machine learning to the characterization of GPS L1 ionospheric amplitude scintillation," 2018 IEEE/ION Position, PLANS, Monterey, CA, 2018, pp. 1159-1166.

[13] X. F. Zhu, X. Y. Chen and J. F. Chen, "Influence of strong ionospheric scintillation on GPS software receiver," Journal of Chinese Inertial Technology, vol. 24, no. 4, pp. 480-484, Aug. 2016.

Detection Method of Solar Radio Bursts Based on Support Vector Machine Model

Yimei Luo
School of Instrument Science and Engineering
Southeast University
Nanjing, China
220193298@seu.edu.cn

Xuefen Zhu*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
zhuxuefen@seu.edu.cn

Mengying Lin
School of Instrument Science and Engineering
Southeast University
Nanjing, China
230198865@seu.edu.cn

Fan Yang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
582919925@qq.com

Abstract—The noise signal produced by solar radio bursts (SRBs) is one of the significant factors influencing the navigation signal. The study of the detection of SRBs is meaningful to the steady operation of navigation system. Taking advantage of the influence of SRBs on Global Navigation Satellite Systems (GNSS) signal, a SRBs detection method based on support vector machine (SVM) model has been proposed. First, the carrier-to-noise ratio, positioning errors of three directions, geometric dilution of precision (GDOP) and the number of satellites loss-of-lock are input to preprocess to obtain the eigenvector and mark whether the SRBs occurs. Then the optimal classifier is obtained by inputting the sample points into SVM classifier for learning. When the new eigenvectors enter the classifier, it will be classified automatically. The average accuracy of SRBs detection reaches more than 93%. This detection method of SRBs based on SVM model can be realized all-time, all-weather and with high efficiency, high accuracy and simple process.

Keywords—GNSS signal, solar radio burst, support vector machine model

I. INTRODUCTION

With the wide application of Global Navigation Satellite Systems (GNSS) and other satellite technologies in modern society, the influence of solar radio bursts (SRBs) intensity on GNSS signal has become an extremely important problem [1]. After Klobuchar's early work, the threat of SRBs to GNSS signals has been widely recognized [2]. Many scholars have conducted extensive research on the interference of some L-band SRBs on GPS in December 2006 [3-5]. Subsequently, Mohamed et al. Reference [6] studied the effect of SRBs on GPS signal, including signal-to-noise ratio, loss-of-lock and phase tracking. Huang et al. Reference [7] selectively analyzed the impact of SRBs event on GNSS performance and positioning error in several typical regions in China on December 13, 2006. A recent study shows that large SRBs are not limited to periods near the maximum sunspot [8].

In recent years, a number of methods of automatic detection and classification of SRBs have been proposed. Ma et al. [9] proposed a method for the classification of SRBs based on multimodal depth learning. The network can effectively learn the representation of solar radio spectrum. Subsequently, Chen et al. Reference [10] achieved better classification results than Ma et al. by utilizing convolutional neural network (CNN) to classify the solar radio spectrum. Singh et al. Reference [11] proposed a method that can distinguish the dynamic spectrum with SRBs and without SRBs automatically.

Most of the methods need to utilize radio telescopes to obtain observation data, and then classify SRBs according to the spectrums. Because radio telescopes are expensive and sparsely distributed. People urgently need a method to detect SRBs in real time and efficiently without utilizing radio telescopes. Zhu et al. Reference [1] proposed a solar radio bursts detection method based on GPS carrier-to-noise ratio decline. Firstly, the solar altitude angle of the observation site is calculated, then the " falling point" and " rising point" are selected to determine the trough time interval of a single satellite in a single observation site. Finally, the detection results of SRBs are obtained by integrating time intervals of multiple satellites and multiple observation sites. The SRBs on December 13, 2006 were detected by the above method. Under the L2 frequency of GPS, when the flux density above 800 SFU , the average accuracy of intense L-band SRBs is more than 80%. However, the test item of this method are single, and can get better results only when the SRBs intensity is strong.

In this paper, a method of SRBs detection based on support vector machine (SVM) model [12—14] is proposed. The test items generated by GNSS signal are comprehensively examined to detect whether the SRBs occur. This method is feasible for the detection of SRBs. The detection results can not only show the occurrence time of SRBs in a single station, but also demonstrate whether multiple stations are affected by SRBs at the same time. Compared with the traditional method, this method does not rely on radio telescope and has low cost. And it combines a variety of factors, the recognition accuracy and efficiency are relatively improved.

II. DETECTION MODEL

This paper analyzes the carrier-to-noise ratio (C/N_0), positioning error, geometric dilution of precision (GDOP) and satellite loss-of-lock of multiple observation sites to detect whether there are SRBs.

A. Preprocess Data

Calculation of the eigenvectors corresponding to the nonoccurrence of SRBs and the occurrence of SRBs at different times in the observation site respectively. The eigenvectors include the C/N_0 reduction, the positioning error, GDOP and the number of satellites loss-of-lock of the observation site.

Firstly, the C/N_0 of each satellite in the observation site, positioning errors of three directions, GDOP and the number

of satellites loss-of-lock are input to data preprocessing.

Take the average of the current C/N_0 reduction value of each satellite as the C/N_0 reduction value of the observation site, and record it as x_1, measured in unit dBHz.

$$x_1 = \sum_{i=1}^{N} snr_i / N \tag{1}$$

Where snr_i is the C/N_0 reduction value of each satellite, that is, average value of C/N_0 of the satellite measured at the first 8 times minus the C/N_0 at current time. N is the number of satellites captured.

The positioning error of the observation site consists of the positioning errors in three directions under the local Cartesian coordinates coordinate system. The standard deviation of the three directions is taken as the positioning error of the observation site, which is recorded as x_2, measured in unit m.

$$x_2 = \sqrt{\frac{\sum_{i=1}^{3} r_i^2}{3}} \tag{2}$$

Where $r_i(i = 1, 2, 3)$ represents the positioning error in different direction of the observation site at current time respectively.

GDOP and the number of satellites loss-of-lock can be expressed by specific numbers, which are recorded as x_3, x_4 respectively.

B. Get Sample Points

The labels corresponding to the eigenvectors of the nonoccurrence of SRBs and the occurrence of SRBs are assigned as -1 and 1 respectively. A set of trained sample points is obtained by combining eigenvectors and labels.

The C/N_0 reduction value of the observation site x_1, the positioning error of the observation site x_2, the GDOP x_3 and the number of satellites loss-of-lock x_4 are combined into the eigenvector x. The sample points are formed as follows:

$$x^{(i)} = \left(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4^{(i)}\right), x^{(i)} \in \mathbb{R}^4 \tag{3}$$

$$y^{(i)} = \begin{cases} -1, & \text{the nonoccurrence of SRB} \\ 1, & \text{the occurrence of SRB} \end{cases} \tag{4}$$

Where $x^{(i)}$ represents the eigenvector of the sample, $y^{(i)}$ represents the label of the sample, the number of the sample points is n, that is, $i = 1, 2, 3, \ldots, n$. And $(x^{(i)}, y^{(i)})$ represents the sample point.

C. Get a Trained Classification Model

To build an unknown nonlinear SVM binary classification model, the discriminant function is defined as

$$g(x) = w^T \Phi(x) + b \tag{5}$$

Where $\Phi(x)$ is a high-dimensional linear mapping function, which maps the 4-dimensional eigenvector x to M-dimensional, that is $\mathbb{R}^4 \rightarrow \mathbb{R}^M$, and $M \gg 4$. For mathematical optimization, the objective of the classification problem is modified to determine $w \in \mathbb{R}^M$, $b \in \mathbb{R}$.

Calculation of the maximum separation hyperplane.

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i \tag{6}$$

$$s.t \begin{cases} y^{(i)} \left(w^T \Phi\left(x^{(i)}\right) + b\right) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases}, i = 1, \ldots, n \tag{7}$$

Where ξ_i is the slack variable of each sample, and the larger ξ_i is, the farther away the sample point is from the group. C is a hyperparameter, and the size of C represents the importance of the loss of objective function caused by outliers. The larger C is, the more attention is paid to outliers, and the tolerance of sample points beyond the maximum boundary is indicated [15].

Solve the Lagrangian multipliers.

$$\max \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y^{(i)} y^{(j)} \tag{8}$$

$$\sum_{i=1}^{n} \alpha_i y^{(i)} = 0, \text{and } 0 \leq \alpha_i \leq C \tag{9}$$

Where α_i, $y^{(i)}$ are Lagrangian multipliers and data classification labels respectively, and $y \in \{-1, 1\}$, n is the number of samples, $i = 1, 2, 3, \ldots, n$.

Furthermore,

$$w_0 = \sum_{i=1}^{n} \alpha_i y^{(i)} \Phi\left(x^{(i)}\right) \tag{10}$$

$$b_0 = y^{(s)} - w_0^{\mathsf{T}} \Phi\left(x^{(s)}\right) \tag{11}$$

where $x^{(s)}$ is the eigenvector of the sample corresponding to the Lagrangian multiplier $\alpha_i \neq 0$, that is, the support vector, and $y^{(s)}$ is the corresponding label.

The calculated w_0 and b_0 are substituted into the expression of nonlinear SVM classifier model.

$$g(x) = \sum_{i=1}^{n} \alpha_i y^{(i)} k\left(x^{(i)}, x\right) + y^{(s)} \\ - \sum_{i=1}^{m} \alpha_i y^{(i)} k\left(x^{(i)}, x^{(s)}\right) \tag{12}$$

where $k(x^{(i)}, x) = \Phi^T(x^{(i)}) \cdot \Phi(x)$, the RBF is selected as the kernel function, which is defined as

$$k\left(x^{(i)}, x\right) = exp\left\{-\frac{\left|\mathbf{x}^{(i)} - \mathbf{x}\right|^2}{2\sigma^2}\right\} \tag{13}$$

Where σ is the kernel parameter.

$$f(x) = sign\left(w^T \Phi\left(x^{(i)}\right) + b\right) \tag{14}$$

When $f(x) = -1$, the SRB does not occur. When $f(x) = 1$, the SRB occurs.

Mark the sample points as $X = (x^{(1)}, x^{(2)}, \ldots, x^{(n)})$, which is a $4 \times n$ matrix. Mark the label as $Y = (y^{(1)}, y^{(2)}, \ldots, y^{(n)})$, which is a $1 \times n$ row vector. The above matrix and vector are

combined into $Z = (X; Y)$ as the sample matrix, which is a $5 \times n$ matrix.

The sample matrix is input into the Classification Learner in MATLAB, the cross validation is chosen and the fold k is set. That is to say, $\frac{k-1}{k} n$ of n samples of matrix Z is randomly selected as training samples for machine learning (k = 2, 3, 4, …), and the remaining $\frac{1}{k} n$ samples are used to test the learned model. The Gaussian nonlinear SVM classifier model is selected, parameters C, $1/(2\sigma^2)$ [16] are set. After training, the average accuracy is obtained.

The values of parameters C and $1/(2\sigma^2)$ are changed, and then a new average accuracy is obtained. Repeat this step until the preset time is reached.

By comparing the average accuracies of all parameters, the hyperparameter C and kernel parameter $1/(2\sigma^2)$ corresponding to the maximum accuracy are found out as the optimal parameters, and the training model with this parameters setting is the optimal classification model.

D. Check Whether SRB Occurs

The eigenvector is extracted from the C/N_0 of each satellite captured, positioning errors of three directions, GDOP and the number of satellites loss-of-lock of the observation site to be detected, which is expressed as $X = (x^{(1)}, x^{(2)}, …, x^{(N)})$, N is the total number of sample points to be detected.

X is input into the classification model, if the eigenvector $x^{(t)}$ makes $f(x^{(t)}) = -1$, the SRB does not occur; otherwise, if $f(x^{(t)}) = 1$, the SRB occurs.

III. Training and Validation

In order to train SVM, the C/N_0 of each captured satellite, positioning errors of three directions, GDOP and the number of satellites loss-of-lock of KUNM, TWTF and XIAN at 1:00:00-3:59:30 (UTC) on December 13, 2006 are selected. Information from IGS data center of Wuhan University(ftp://igs.gnsswhu.cn/).

The sampling interval is 30s, and each observation site has 360 times in total.

A. Detection Experiment Process

The flow chart of SRBs detection is shown in Fig.1.

Firstly, the data are preprocessed, the C/N_0 reduction value, the positioning error, the GDOP and the number of satellites loss-of-lock of the observation site are taken as eigenvectors. Based on the fact that SVM is a binary classifier, the occurrence of SRB is marked as 1, the nonoccurrence is marked as -1. Then labels and eigenvectors are combined into sample points. To obtain a trained classification model, an unknown nonlinear SVM binary classification model is built, and the optimal parameters are obtained by cross validation of training samples and comparison of average accuracies.

A. Single Observation Site Detection Process

Take KUNM as an example, after data preprocessing and label allocation, the sample points of KUNM at different times are obtained. Table I lists 5 relatively representative sample points. Other sample points are processed in the same way, which are not described in detail.
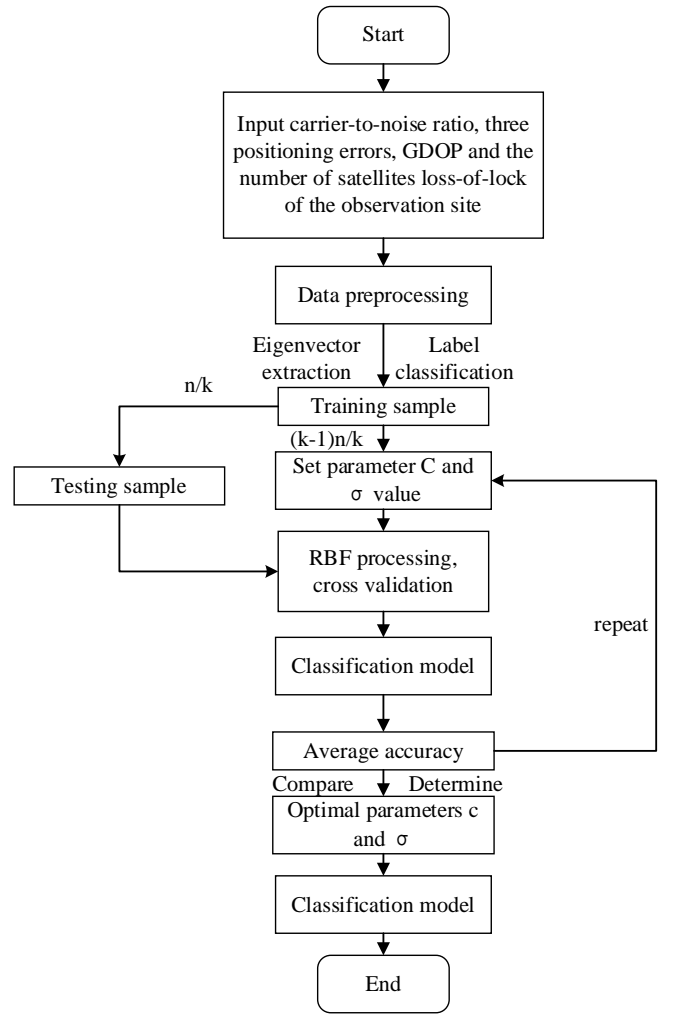
Fig.1 Flow chart of SRBs detection.

TABLE I. SAMPLE POINTS CORRESPONDING TO DIFFERENT TIME OF OBSERVATION SITE KUNM

Time	Sample point
2:00:00	(-0.40625, 3.82452, 1.8,0,-1)
2:59:30	(-0.47222, 9.86594, 2.2,1,1)
3:31:30	(7.67708, 11.80417, 16.2,4,1)
3:36:00	(11.44792, /, /,5,1)
3:52:00	(-0.06548, 7.79374,2.1,1,-1)

In Table I, " / " represents unable to get position or GDOP.

The 360 sample points are input into matrix Z, the cross validation fold is set to $k = 5$, the parameters are changed.

$$C = 2^{-1}, 1, 2, 2^2, 2^3, 2^4, 2^5, …$$

$$1/(2\sigma^2) = 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, …$$

Then, the average accuracy with different parameters is obtained by comparison.

When $C = 2^3 = 8$, $1/(2\sigma^2) = 0.5$, the classifier has the highest average accuracy of 94.7%, so the training model with

this parameters setting is the optimal classification model. The confusion function with the optimal parameters setting is as shown in the Fig. 2.
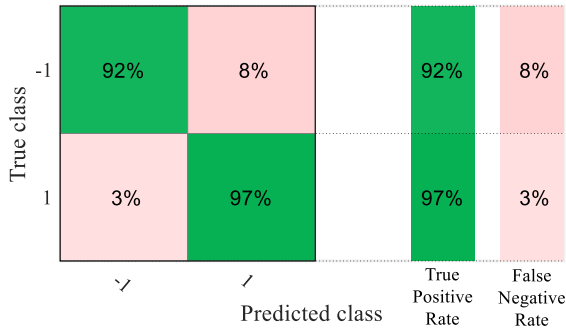
Fig.2 Confusion matrix of SRBs prediction results in KUNM.

Each column of the confusion matrix represents the forecast category, each row represents the real category of data. The larger the total percentage of the matrix diagonal, the better the classifier performs.

As shown in the Fig. 2, 92% of the sample points without SRBs and 97% of the sample points with SRBs are classified correctly.

B. Detection Results of Multiple Observation Sites

In TWTF, when $C = 8$, $1/(2\sigma^2) = 8$, the classifier has the highest average accuracy of 96.7%. The confusion function with the optimal parameters setting is shown in Fig. 3a.

In XIAN, when $C = 64$, $1/(2\sigma^2) = 2$, the classifier has the highest average accuracy of 95.3%. The confusion function with the optimal parameters setting is shown in Fig. 3b.

Fig.3 (a) Confusion matrix of SRBs prediction results in TWTF. (b) Confusion matrix of SRBs prediction results in XIAN.

IV. CONCLUSION

Based on the SVM model, a new method for detecting SRBs is proposed. SVM learns the distinguishing features from the training data to maximize the margin of separation between the two classes. The method makes a comprehensive examination of the tested items generated by GNSS signals, and comprehensively judges whether SRBs occur by detecting the C/N_0, positioning error, GDOP and the number of satellites loss-of-lock of the observation site.

A large amount of different observation sites data collected facilitated the training and testing of the SVM detector. Through several experiments, the proposed detection algorithm is proved high reliability. The overall accuracy is above 93% for historical SRB events.

Compared with the traditional method, the major benefit of this method is that it does not require the radio telescope, and can achieve all-time, all-weather detection. The method processes large quantities of data at the same time and the detection results demonstrate whether multiple stations are affected by SRBs. It not only saves cost, but also improves the detection efficiency and accuracy. The proposed method and the relevant experience in this paper can be conducive to maintain the normal operation of the satellite navigation system.

REFERENCES

[1] X. F. Zhu, F. Yang, X. Y. Chen, M. Y. Lin, and X. Chen. Detection method of solar radio bursts based on the decrease of GPS carrier to noise ratio[J/OL]. Systems Engineering and Electronics, 2020, pp. 1-8.

[2] J. A. Klobuchar, J. M. Kunches, and A. J. VanDierendonck. Eye on the Ionosphere: Potential Solar Radio Burst Effects on GPS Signal to Noise[J]. GPS Solutions, 1999, 3(2): pp. 69–71.

[3] C.S. Carrano, C. T. Bridgwood, and K. M. Groves. Impacts of the December 2006 solar radio bursts on the performance of GPS[J]. Radio Science, 2009, 44(1): pp.1-12.

[4] A. P. Cerruti, P. M. Kintner, D. E. Gary, A. J. Mannucci, R. F. Meyer, P. Doherty, and A. J. Coster. Effect of intense December 2006 solar radio bursts on GPS receivers[J]. Space Weather, 2008, 6(10): pp.1-10.

[5] P. M. Kintner, B. O'Hanlon, D. E. Gary, and P. M. S. Kintner. Global Positioning System and solar radio burst forensics[J]. Radio Science, 2009, 44(1): pp.1-6.

[6] B. Muhammad, V. Alberti, A. Valentina, A. Marassi, and C. Ernestina, Performance assessment of GPS receivers during the September 24, 2011 solar radio burst event[J]. Journal of Space Weather & Space Climate, 2015, 5: pp. 1-16.

[7] W. G. Huang, A. Ercha, H. Shen, S. Q. Liu, and L. Q. Shi. Impact of intense L-band solar radio burst on GNSS performance and positioning accuracy[J]. Chinese Journal of Radio ence, 2018, 33(1): pp. 1-7.

[8] O. D. Giersch, J. Kennewell, and M. Lynch. Solar radio burst statistics and implications for space weather effects[J]. Space Weather, 2017, 15: pp. 1511–1522.

[9] L. Ma, Z. Chen, L. Xu, and Y. H. Yan. Multimodal deep learning for solar radio burst classification[J]. Pattern Recognition, 2016, 61: pp. 573-582.

[10] S. Chen, L. Xu, L. Ma, and W. Q. Zhang. Convolutional neural network for classification of solar radio spectrum[C]. 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, 2017, pp. 198-201.

[11] D. Singh, K. S. Raja, P. Subramanian, R. Ramesh, and C. Monstein. Automated Detection of Solar Radio Bursts using a Statistical Method[J]. Solar Physics, 2019, 249(8): pp. 1-14.

[12] X. B. Gao and J. P. Zhang. Machine learning and its application 2015 [M]. Tsinghua University Press, 2015.

[13] Haykin, S. Neural networks and learning machines[M], China Machine Press, 2009, pp. 268-312.

[14] L. C. Xiong and Y. M. Li. Five understandings of objective function in support vector machine regression model[J]. Hans Journal of Data Mining, 2019, 9 (2): pp.52-59.

[15] L. M. Liang, Z. Zhong, and Z. Y. Chen. Research and Simulation of Kernel function selection for support vector machine[J]. Computer Engineering and Science, 2015, 37 (6): pp.1135-1141.

[16] G. H. Feng. Parameter optimizing for Support Vector Machines classification[J]. Computer Engineering and Applications , 2011, 47(3): pp. 123-124.

Optimal Placement of Blade Tip Timing Sensors Considering Multi-mode Vibration Using Evolutionary Algorithms

Jinghui Xu
The State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
denniss@stu.xjtu.edu.cn

Baijie Qiao*
The State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
qiao1224@xjtu.edu.cn

Zhibo Yang
The State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
phdapple@mail.xjtu.edu.cn

Yuanchang Chen
College of Engineering
University of Massachusetts Lowell
Massachusetts, United States
y.c.chen1990@gmail.com

Xuefeng Chen
The State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
chenxf@mail.xjtu.edu.cn

Abstract—Blade tip timing (BTT) as a noncontact measurement technique is developed to monitor the health condition of aero-engine rotor blades. The circumferential placement of BTT sensors plays an important part in identifying blade vibration parameters. If the placement is suboptimal, the reconstructed displacement signal can be sensitive to the measurement noise while certain useful information is lost. This paper presents an optimal placement of blade tip timing sensors considering multi-mode vibration using evolutionary algorithms. The installation angle of BTT sensors around the casing is taken as the design variable and the inaccessible angle of the casing is taken as the constraint condition. Particle swarm optimization (PSO) algorithm is the preferred optimization one in this work due to its strong robustness and fast convergence. The condition number of the design matrix of blade multi-mode vibration reconstruction model is taken as the fitness function of PSO, the minimum of which corresponds to the optimal sensor positions. Simulation experiment results show that compared with the random placement of BTT sensors, the more blade modes required to be identified along with the circumferential Fourier fit algorithm employed, the more prominent the advantage of the proposed method is.

Keywords—Blade tip timing, sensor optimal placement, evolutionary algorithms, aero-engine blades

I. INTRODUCTION

Rotating blades are important parts of aero-engine. The harsh operation environment of aero-engine makes the blades vibrate easily, which can lead to the high cycle fatigue (HCF) crack and other damages [1]. Airlines invest huge amounts of money in engine maintenance and monitoring every year. Condition based maintenance is pursued by current airlines for ensuring flight safety and reducing the cost of operation and maintenance of aero-engine [2]. The vibration state of the blades could be changed due to the fatigue cracks and other damages [3]. Therefore, it is necessary to monitor the vibration parameters of blades in real time to know the operation condition and detect the damage of blades in the early period after damage growth. Online monitoring provides sufficient data for the evaluation of operation condition of aero-engine and is significant to reduce the cost of maintenance and ensure the safety of operation.

Traditionally, the strain gage gluing on the surface of the blade is commonly used to measure the dynamic stress of the rotor blades. However, such a contact measurement technique is limited to finite blades and then the strain gage often fails to record useful data in harsh operating condition. The noncontact method should be developed to monitor the operation condition of the blades efficiently. Recently, blade tip timing (BTT) as a noncontact measurement technique that can measure all the blade of the same stage has been developed to monitor the health condition of aero-engine rotor blades. BTT sensors are installed on the casing to detect the arrival time of the rotor blades. The vibration displacement of blade tip can be derived from sensor data and further vibration parameters including vibration frequency, amplitude and phase can be identified by specific algorithms.

However, the data collected by BTT sensors is under-sampled because a blade can pass through a BTT sensor only one time in a revolution. In other words, the sampling frequency equals to the rotation frequency, if it has only one BTT sensor. Frequency aliasing can result from traditional spectrum analysis methods in Nyquist sampling theorem. In this case, it is difficult to identify blade vibration parameters accurately [4]. Although there exist some data processing algorithms working well for BTT signals [5], if the placement is suboptimal, the reconstructed displacement signal can be sensitive to measurement noise while certain useful information is lost [6]. Diamond et al. [6]proposed a BTT sensor placement method, but only considered a single vibration mode of blade. The actual vibration state of aero-engine blades may be a multi-mode one, which means multiple natural frequencies of blades are simultaneously excited during the operation.

This paper presents a placement method of BTT sensors for identifying the multi-modal vibration of blades using evolutionary algorithms. It reduces the identification errors of multi-mode vibration parameters of blades using circumferential Fourier fit (CFF) and improve the evaluation accuracy of operation state of blades.

II. BLADE TIP TIMING TECHNIQUE

The radial vibration of the blades can be controlled by the high precision dynamic balance of the rotor. The most

---

dangerous vibration of rotating blade is the circumferential vibration, which may lead to reduce the HCF life. Therefore, the circumferential vibration displacement of the blades in the tangential direction can be monitored by BTT system. The schematic diagram of BTT system is shown in Fig. 1.

Sensors are installed in the radial direction on the casing of aero-engine. When the blades pass through the BTT sensors, the reflected intensity of the signal increases suddenly and the arrival time of the blades can be accurately measured by setting the threshold of signal reflection intensity. If the blades operate without vibration, the arrival time of blades is expressed in the form of equally spaced pulses as shown in Fig. 2.

The vibration of the blades in the circumferential direction changes the arrival time of blades as shown in Fig. 3. If the blades arrive early, it means the amplitude of the blades vibration is positive. The lag of the blades means the negative amplitude of the blades vibration.

The vibration displacement of blades d can be calculated according to the rotation frequency f_n, blade radius R and time difference Δt.

$$d = v\Delta t = 2\pi f_n R \Delta t \tag{1}$$

When the blades rotate s cycles, the BTT system can collect $s \times n$ points of vibration displacement, where n is the number of the BTT sensors. It means the sampling frequency f_s equals to $n \times f_\omega$, where f_ω is the rotation frequency. In most cases, the natural frequency of the blades of interest is more than $2 \times n \times f_\omega$. In most situations, it is difficult for sampling data to satisfy the theorem of Nyquist sampling. Therefore, some special algorithms are developed to process BTT signals.
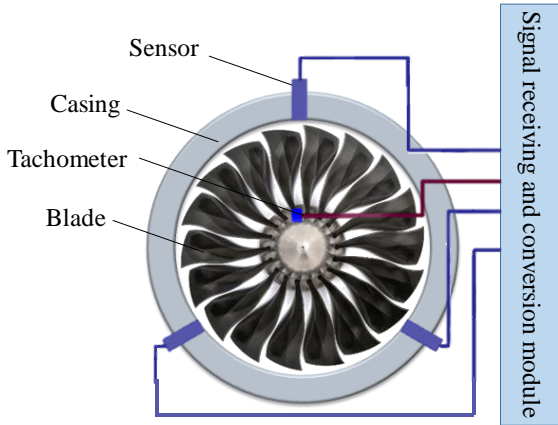
Fig. 1 Schematic diagram of BTT measuring the vibration of the rotor blade
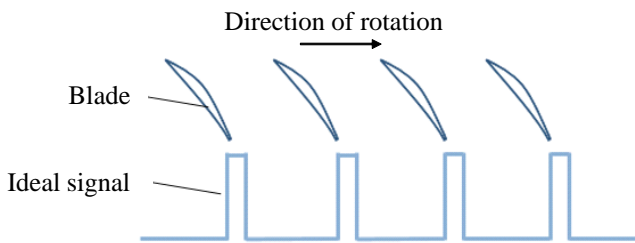
Fig. 2 Measuring pulses of the blade tip without vibration
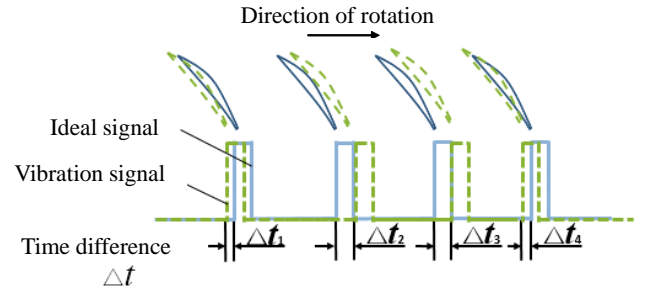
Fig. 3 Measuring pulses of the blade tip with vibration

III. MATHEMATICAL MODEL OF BLADE VIBRATION

The blade has infinite degrees of freedom in theory as continuums. A continuous blade is discretized by finite element method, and only the first m order vibration modes are considered. Dimitriadis et al. [7] stated that if only a single vibration mode is considered, the vibration equation of the blade can be expressed as follows:

$$y = A\sin(2\pi ft) + B\cos(2\pi ft) + C \tag{2}$$

where A, B and C are the parameters with respect to the amplitude, phase and offset of the vibration response. f is the vibration frequency of the blade. When the synchronous vibration of the blade occurs, the relationship between the vibration frequency of the blade and the rotation frequency is:

$$f = EO \times f_\omega \tag{3}$$

where f_ω is the rotation frequency of the blade and EO denotes the engine order. Furthermore, Eq. (2) can be rewritten as

$$y = A\sin(2\pi \times EO \times f_\omega t) + B\cos(2\pi \times EO \times f_\omega t) + C \tag{4}$$

The rotation phase angle is defined as:

$$\theta_i = 2\pi f_\omega t \tag{5}$$

Substituting Eq. (5) into Eq. (4):

$$y = A\sin(EO\theta_i) + B\cos(EO\theta_i) + C \tag{6}$$

If the first m order vibration modes of blades are simultaneously considered, the natural frequencies of first m order are $f_1, f_2, ..., f_m$. Then the engine orders of the blade are given in the following equation:

$$EO = \left\{\frac{f_\omega}{f_1}, \frac{f_\omega}{f_2}, ..., \frac{f_\omega}{f_m}\right\} = \left\{EO_1, EO_2, ..., EO_m\right\} \tag{7}$$

Therefore, the displacement of certain a blade measured by j th BTT sensor can be analytically expressed：

$$y_j = \sum_{i=1}^{m}[A_i \sin(EO_i\theta_j) + B_i \cos(EO_i\theta_j) + C_i] \tag{8}$$

Furthermore, the blade displacement measured by n BTT sensors can be expressed in a matrix-vector form:

$$\mathbf{y} = \mathbf{Hx} \tag{9}$$

where $\mathbf{H}$ is called the design matrix, defined as:

$$\mathbf{H} = \begin{pmatrix} \sin(EO_1\theta_1) & \cos(EO_1\theta_1) & \sin(EO_2\theta_1) & \cos(EO_2\theta_1) & \cdots & 1 \\ \sin(EO_1\theta_2) & \cos(EO_1\theta_2) & \sin(EO_2\theta_2) & \cos(EO_2\theta_2) & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sin(EO_1\theta_n) & \cos(EO_1\theta_n) & \sin(EO_2\theta_n) & \cos(EO_2\theta_n) & \cdots & 1 \end{pmatrix} \tag{10}$$

$\mathbf{x}^{n\times 1} = \begin{pmatrix} A_1 & B_1 & A_2 & B_2 & \dots & C \end{pmatrix}^{\mathrm{T}}$ is an unknown vector of multi-mode vibration parameters. $\mathbf{y}^{n\times 1} = \begin{pmatrix} y_1 & y_2 & \dots & y_n \end{pmatrix}^{\mathrm{T}}$ is the displacement vector measured by n BTT sensors in the same revolution. $\mathbf{x}$ can be calculated by CFF

$$\mathbf{x} = \mathbf{H}^{-1}\mathbf{y} \tag{11}$$

The number of sensors n should satisfy Eq. (12) to ensure Eq. (11) having a solution.

$$n \geq 2m+1 \tag{12}$$

where m is the number of EOs.

Such a procedure including Eqs. (8)-(12) is namely the circumferential Fourier fit method, assuming the rotation speed is constant in a revolution. Order tracking method is used by CFF to identify the amplitude and phase of the vibration [8]. CFF looks at one blade from all sensors. Therefore, the signal is a representation of the behavior of one blade through a revolution. The condition number of $\mathbf{H}$ gives the ratio of the maximum and minimum singular values. Then, the condition number of $\mathbf{H}$ is defined as:

$$\mathrm{cond}(\mathbf{H}) = \|\mathbf{H}\| \cdot \|\mathbf{H}^{-1}\| \tag{13}$$

The small condition number means that the placement of BTT sensors is orthogonal approximately. In other words, there exists little information redundancy. The smaller the condition number of $\mathbf{H}$ is, the less sensitive CFF is to noise in measurement.

IV. Particle Swarm Optimization

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

Particle swarm optimization (PSO) is preferred in this paper due to its efficiency and convenience. The condition number of the design matrix is taken as the fitness function of PSO. The installation angle of BTT sensors is taken as the design variable and the inaccessible angle of BTT sensors is taken as the constraint condition. The PSO procedure for BTT sensors optimal placement includes the following steps:

Step 1: Initialization: N groups of the installation angle of n BTT sensors are randomly generated. It should generate a random matrix $\boldsymbol{\Psi}_0$ of N rows and n rows within the constraints of the circumferential accessible angle on the casing. $\boldsymbol{\Psi}_0$ represents the placement of N groups of BTT sensors at the initial state, when iteration $iter = 0$. The velocity of PSO $(\mathbf{V}_0)_{N\times n}$ is randomly initialized within the constraint range.

Step 2: The condition number of the design matrix for N groups of BTT sensors is calculated separately. The minimum value of the condition number is recorded as $\kappa_{\min}$ and the placement of BTT sensors corresponding to $\kappa_{\min}$ is recorded as $(\boldsymbol{\theta}_{\min})_{1\times n}$. The placement of N groups of BTT sensors is recorded as $\boldsymbol{\Theta}_{N\times n}$, which is the initialization of the historical optimum installation angle for each group of BTT sensors.

Step 3: The velocity of PSO $\mathbf{V}_{iter+1}$ is updated:

$$\begin{aligned} \mathbf{V}_{iter+1} = {}& \omega\mathbf{V}_{iter} + C_1\mathrm{rand}(0,1)(\boldsymbol{\Theta}_{iter} - \boldsymbol{\Psi}_{iter}) \\ & + C_2\mathrm{rand}(0,1)(\mathrm{repmat}(\boldsymbol{\theta}_{\min}, N) - \boldsymbol{\Psi}_{iter}) \end{aligned} \tag{14}$$

where ω is the inertia factor, C_1 and C_2 are the constants, $\mathrm{rand}(0,1)$ represents random numbers on intervals $(0, 1)$, $\mathrm{repmat}(\boldsymbol{\theta}_{\min}, N)$ represents copying a group of optimal placement of BTT sensors $\boldsymbol{\theta}_{\min}$ to N dimensions for guaranteeing the additivity of the matrix.

Step 4: The design variable of PSO $\boldsymbol{\Psi}_{iter+1}$ is updated:

$$\boldsymbol{\Psi}_{iter+1} = \boldsymbol{\Psi}_{iter} + \mathbf{V}_{iter+1} \tag{15}$$

Step 5: The history of optimal installation angle for each group of BTT sensors $\boldsymbol{\Theta}$, optimal installation angle for all groups of BTT sensors $\boldsymbol{\theta}_{\min}$ and the minimum condition number $\kappa_{\min}$ are updated. When the accuracy of fitness function does not reach the set value, repeat steps 3, 4 and 5 and $iter = iter+1$. After the iteration, $\boldsymbol{\theta}_{\min}$ is the installation angle of the BTT sensors, which is helpful in identifying the multi-mode vibration parameters.

V. Illustrative Example

A non-torsion blade is established to verify the proposed method as shown in Fig. 4. The blade is 48 mm long, 20 mm wide and 1 mm thick. First three natural frequencies are considered. The first three natural frequencies of blades are listed in Table I.

Fig. 4 A blade model

When the vibration frequency of the blade is close to the natural frequency, the amplitude of the vibration increases significantly and the blade is inclined to induce crack. Therefore, it requires to focus on EO which may excite the natural frequencies of the blade. Here, assuming the rotation frequency of the blade is 100 Hz, the EOs required to be paid attention as follows:

TABLE I.　FIRST THREE NATURAL FREQUENCIES OF THE BLADE

Modal	Natural frequency
Mode 1: 1st bending	333.08 Hz
Mode 2: 1st Torsion	1806.03 Hz
Mode 3: 2nd bending	2076.52 Hz

$$EO_1 = \frac{333.08}{100} \approx 3, \ EO_2 = \frac{1806.03}{100} \approx 18, \ EO_3 = \frac{2076.52}{100} \approx 21 \quad (16)$$

According to Eq. (8), the blade vibration equation involving three modes at the same time is written by the following equation:

$$y_i = A_1 \sin(EO_1\theta_i) + B_1 \cos(EO_1\theta_i) + A_2 \sin(EO_2\theta_i)$$
$$+ B_2 \cos(EO_2\theta_i) + A_3 \sin(EO_3\theta_i) + B_3 \cos(EO_3\theta_i) + C \quad (17)$$

where y_i is the displacement measured by the i th BTT sensor, θ_i is the installation angle of the i th BTT sensor. There are seven unknown parameters in Eq. (17). Seven BTT sensors are required to measure the displacement at least. Installation angle of the BTT sensors is limited to $U = [0, 360°)$ and the matrix $\Psi \in U^{300 \times 7}$ of placement of BTT sensors is generated randomly. The design matrix for each group of BTT sensors is calculated. Take the first group of BTT sensors as an example, the initial installation angle is generated randomly:

$$\theta = (12.1°, \ 76.8, \ 83.7°, \ 117.5, \ 189.6°, \ 303.7°, \ 339.7°)$$

Then the corresponding design matrix is:

$$H = \begin{pmatrix} \sin(3 \times 0.21) & \cos(3 \times 0.21) & \sin(18 \times 0.21) & \cos(18 \times 0.21) & \sin(21 \times 0.21) & \cos(21 \times 0.21) & 1 \\ \sin(3 \times 1.34) & \cos(3 \times 1.34) & \sin(18 \times 1.34) & \cos(18 \times 1.34) & \sin(21 \times 1.34) & \cos(21 \times 1.34) & 1 \\ \sin(3 \times 1.46) & \cos(3 \times 1.46) & \sin(18 \times 1.46) & \cos(18 \times 1.46) & \sin(21 \times 1.46) & \cos(21 \times 1.46) & 1 \\ \sin(3 \times 2.05) & \cos(3 \times 2.05) & \sin(18 \times 2.05) & \cos(18 \times 2.05) & \sin(21 \times 2.05) & \cos(21 \times 2.05) & 1 \\ \sin(3 \times 3.31) & \cos(3 \times 3.31) & \sin(18 \times 3.31) & \cos(18 \times 3.31) & \sin(21 \times 3.31) & \cos(21 \times 3.31) & 1 \\ \sin(3 \times 5.30) & \cos(3 \times 5.30) & \sin(18 \times 5.30) & \cos(18 \times 5.30) & \sin(21 \times 5.30) & \cos(21 \times 5.30) & 1 \\ \sin(3 \times 5.93) & \cos(3 \times 5.93) & \sin(18 \times 5.93) & \cos(18 \times 5.93) & \sin(21 \times 5.93) & \cos(21 \times 5.93) & 1 \end{pmatrix}$$

The condition number of H calculated by Eq. (13) is 46.37.

The condition number of all groups of BTT sensors can be calculated by Eq. (13). And the result is defined as a vector $\varphi \in R^{300 \times 1}$.

After the calculation, the optimal placement of BTT sensors is:

$$\theta_{min} = (31.7°, \ 56.6°, \ 89.7°, \ 120.4°, \ 146.1°, \ 184.9°, \ 215.2°)$$
And the placement of BTT sensors is shown in Fig. 5.
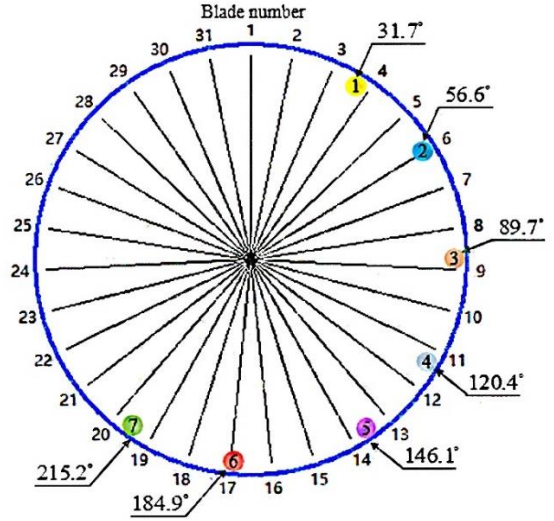
Fig. 5 The optimal placement of BTT sensors

CFF is used to verify the reconstruction performance of blade vibration parameters. The basic principle is expressed in Eq. (9). According to the measurements of seven BTT sensors and the design matrix of the placement of BTT sensors, the vibration parameters can be calculated by Eq. (11).

The rotation frequency of the blades is $f_\omega = 100$Hz and the vibration parameters in Eq. (17) is:

$$A_1 = 85; B_1 = 73; A_2 = 54; B_2 = 32; A_3 = 28; B_3 = 19; C = 12$$

The Gaussian white noise is used to simulate the uncertainty in the measurement. Measurement uncertainty is expressed by signal-to-noise ratio:

$$SNR = 10 \times \lg \frac{P_{signal}}{P_{noise}} \quad (18)$$

where P_{signal} and P_{noise} present the effective power of signal and noise.

When SNR=5 dB, data points collected by optimal placement of BTT sensors and reconstructed signals by using CFF are shown in Fig. 6. Fig. 6 illustrates that the reconstructed signal matches well with the original signal. Although some measurement points deviate from the original signal, the reconstructed results are not greatly affected. Because the condition number of the design matrix is small, the variation of measurements has little influence on the solution of vibration parameters. A randomly generated placement of BTT sensors is used as a comparison. 1,000 groups of placement of BTT sensors are generated randomly. A group with the minimum condition number is chosen as a control group.

The installation angle of BTT sensors of the control group is:

$$\theta = \begin{pmatrix} 85.2° & 132.2° & 152.2° & 200.5° & 262.7° & 285.8° & 312.4° \end{pmatrix}$$

The corresponding BTT sensors placement is shown in Fig. 7.

Fig. 6 Comparing reconstructed signal using BTT sensors optimal placement with the original signal and noisy signal in a revolution

Fig. 7 The random placement of BTT sensors

The condition number of optimal placement of BTT sensors is 1.71 while the condition number of control group is 3.83. Data points collected by random placement of BTT sensors and reconstructed signals by using CFF are shown in Fig. 8.

Fig. 8 Comparing reconstructed signal using BTT sensors random placement with the original signal and noisy signal in a revolution
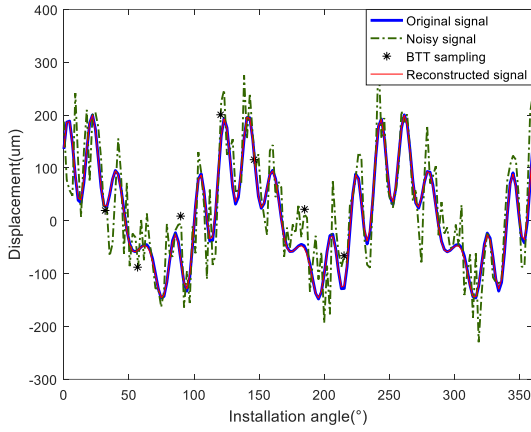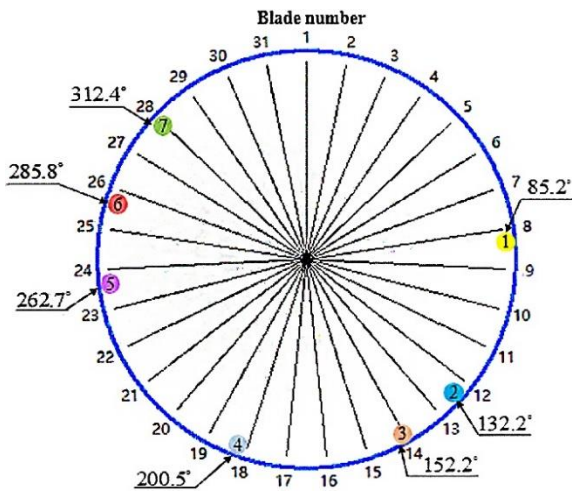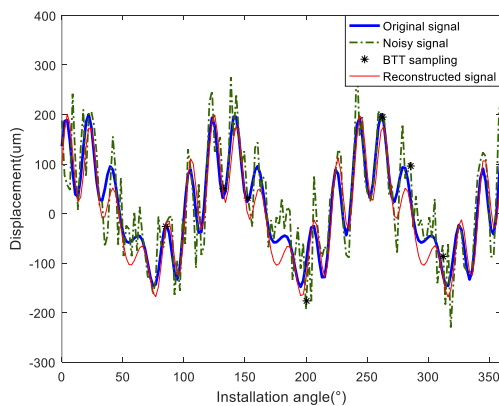
Fig. 8 illustrates that the reconstructed signal has large errors compared with original signal particularly near the peak points. The measurements of BTT sensors are closer to the original signal compared with that in Fig. 6. However, the

condition number of the design matrix is larger than optimal placement. It leads to the poor robustness. The reconstructed signal is easier to be influenced if the measurements have a small change. If the placement is suboptimal, the reconstructed signal is sensitive to measurement noise.

Furthermore, relative errors of vibration parameters are quantitatively used to express the difference between two different schemes of placement of BTT sensors.

The equation for calculating relative errors is given as:

$$\delta = \frac{\sqrt{\sum_{i=1}^{3}[(A_i - \tilde{A}_i)^2 + (B_i - \tilde{B}_i)^2] + (C - \tilde{C})^2}}{\sqrt{\sum_{i=1}^{3}(A_i^2 + B_i^2) + C^2}} \times 100\% \quad (19)$$

where A_i, B_i and C are the parameters set in Eq. (17), and $\tilde{A}_i$, $\tilde{B}_i$ and $\tilde{C}$ are the vibration parameters calculated by CFF. The case runs 1,000 times repeatedly with the different noise. The averaged relative error of vibration parameters of the optimal placement is $\delta_1 = 15.1\%$ and the averaged relative error of vibration parameters of the random placement is $\delta_2 = 21.8\%$.

As for the computing time, under the condition of identifying three modes, it takes 0.65s with 2.2GHz CPU and 8G memory to calculate the optimal placement of the BTT sensors by using MATLAB. If the placement of BTT sensors is randomly generated, it requires to generate 500, 000 times to obtain the same condition number of the design matrix and it only consumes 180s. Meanwhile, the more blade modes required to be identified along with the circumferential Fourier fit algorithm employed, the more prominent the advantage of the proposed method is.

To verify the effectiveness of the method, the optimal placement of BTT sensors are compared with other scholars used such as "5+2" method, which could eliminate some aliased components of synchronous vibration [9]. According to "5+2" method, installation angles of seven BTT sensors are set as:

$$\mathbf{\theta}_{5+2} = \begin{pmatrix} 0° & 72° & 120° & 144° & 216° & 240° & 288° \end{pmatrix}$$

However, the condition number of "5+2" placement is 6.05×10^{16}, which means that the design matrix is a singular matrix. Therefore, Eq. (11) should be extended as:

$$\mathbf{x} = \mathbf{H}^\dagger \mathbf{y} \quad (20)$$

where $\mathbf{H}^\dagger$ is the Moore-Penrose pseudoinverse. Similarly, "5+2" placement is used to reconstruct the simulation signal generated by Eq.(17). 1000 trials are tested under the signal-to-noise ratio of 5dB. One of the reconstructed result is illustrated in Fig. 9. The averaged relative error of vibration parameters of the optimal placement is $\delta_3 = 30.12\%$.

Comparing Fig. 6, Fig. 8 and Fig. 9, it can be inferred that it is useful to improve the accuracy of reconstructed signal by reducing the condition number of design matrix. Although the signal can be reconstructed by Moore-Penrose pseudoinverse using "5+2" placement, the accuracy is lower than the optimal placement and random placement.
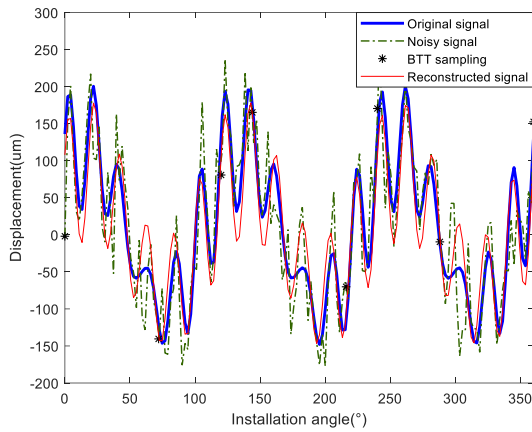
Fig. 9 Comparing reconstructed signal using BTT sensors random placement with the original signal and noisy signal in a revolution

REFERENCES

[1] G. Battiato, C. M. Firrone and T. M. Berruti, "Forced response of rotating bladed disks: Blade Tip-Timing measurements," *MECHANICAL SYSTEMS AND SIGNAL PROCESSING,* vol. 85, pp. 912-926, 2017.

[2] X. Rong and H. Zuo, "Condition-based maintenance decision making methods for civil aeroengine," *Transactions of Nanjing University of Aeronautics & Astronautics,* vol. 25, pp. 201-7, 2008.

[3] Z. Chen, Y. Yang, Y. Xie, B. Guo, and Z. Hu, "Non-contact crack detection of high-speed blades based on principal component analysis and Euclidian angles using optical-fiber sensors," *SENSORS AND ACTUATORS A-PHYSICAL,* vol. 201, pp. 66-72, 2013.

[4] Z. Hu, J. Lin, Z. Chen, Y. Yang, and X. Li, "A Non-Uniformly Under-Sampled Blade Tip-Timing Signal Reconstruction Method for Blade Vibration Monitoring," *SENSORS,* vol. 15, pp. 2419-2437, 2015.

[5] S. Heath and M. Imregun, "A Review of Analysis Techniques for Blade Tip-Timing Measurements," in *ASME 1997 International Gas Turbine and Aeroengine Congress and Exhibition.* vol. 4 Orlando, Florida, USA, 1997, pp. V004T15A016.

[6] M. Mohamed, P. Bonello and P. Russhard, "A novel method for the determination of the change in blade tip timing probe sensing position due to steady movements," *Mechanical Systems and Signal Processing,* vol. 126, pp. 686-710, 2019.

[7] G. Dimitriadis, I. B. Carrington, J. R. Wright, and J. E. Cooper, "Blade-tip timing measurement of synchronous vibrations of rotating bladed assemblies," *MECHANICAL SYSTEMS AND SIGNAL PROCESSING,* vol. 16, pp. 599-622, 2002.

[8] D. H. Diamond, P. S. Heyns and A. J. Oberholster, "A Comparison Between Three Blade Tip Timing Algorithms for Estimating Synchronous Turbomachine Blade Vibration," in *Lecture Notes in Mechanical Engineering*, J. AmadiEchendu, C. Hoohlo and J. Mathew, Eds., 2015, pp. 215-225.

[9] Beauseroy P , Lengelle R. "Nonintrusive turbomachine blade vibration measurement system," *MECHANICAL SYSTEMS AND SIGNAL PROCESSING*, vol. 21, pp. 1717-1738, 2007

Eddy Current Thermography for the Detection of Conductive Defects in Composite Insulators

Yanxin Tu
Engineering Laboratory of Power Equipment Reliability in Complicated Coastal Environments
Tsinghua Shenzhen International Graduate school
Shenzhen, China
tyx18@mails.tsinghua.edu.cn

Chenjun Guo
Electric Power Research Institute
Yunnan Power Grid Co., Ltd
China Southern Power Grid Co., Ltd
Kunming, China
807668414@qq.com

Hongwei Mei
Engineering Laboratory of Power Equipment Reliability in Complicated Coastal Environments
Tsinghua Shenzhen International Graduate school
Shenzhen, China
mei.hongwei@sz.tsinghua.edu.cn

Lishuai Liu
Engineering Laboratory of Power Equipment Reliability in Complicated Coastal Environments
Tsinghua Shenzhen International Graduate school
Shenzhen, China
lls17@mails.tsinghua.edu.cn

Chenglong Cong
Engineering Laboratory of Power Equipment Reliability in Complicated Coastal Environments
Tsinghua Shenzhen International Graduate school
Shenzhen, China
congcl19@mails.tsinghua.edu.cn

Liming Wang*
Engineering Laboratory of Power Equipment Reliability in Complicated Coastal Environments
Tsinghua Shenzhen International Graduate school
Shenzhen, China
wanglm@sz.tsinghua.edu.cn

Abstract—**Composite insulators are important and fundamental electrical equipment in power systems. However, the internal defects caused by long-term outdoor operation also cause great damage to the power system. For the purpose of eliminating the internal insulation breakdown which is caused by internal conductive defects of composite insulator, this work proposed a method for detecting internal conductive defects of composite insulator based on Eddy current thermography (ECT). This method uses the external excitation source to heat the conductive defects, internal defects can be detected through processing and analyzing the thermal image sequence captured by thermal camera. The effectiveness of this method is verified by analyzing original thermal image sequence. The detection effect of defects has been enhanced by using Principal component analysis (PCA) and independent component analysis (ICA) methods. The theoretical and experimental results indicate that Eddy current thermography has a broad application prospect in the detection testing of conductive flaws in composite insulators.**

Keywords—composite insulator, conductive defects, Eddy current thermography, principal component analysis, independent component analysis

I. INTRODUCTION

As an important element of power transmission system, composite insulator has been widely employed for their low weight, high mechanical strength, strong anti-pollution flashover ability[1]. However, with the increase of service time, composite insulators will be affected by the external environment, such as high electric field, high humidity, and high temperature. Due to the particularity of on-site operation, composite insulators are prone to generate internal defects, which may lead to insulation failure and even power interruption[2]. Some of the research from on-site composite insulators shows that, there are hidden defects in the interior of the composite insulator after a long period of operation, causing insulation failure of the composite insulator. Insulation failure induced by the inner conduction of composite insulator is one of the severe faults of composite insulator, which may threaten the normal running of power delivery system[3]. As shown in Fig. 1, the core rod of the composite insulator from power lines are hydrolyzed under the erosion of electricity, and the epoxy resin core rod is in the shape of dry wood, which may cause the composite insulator to break. Recently, several composite insulator rupture accidents may be caused by the development of inner conductive defects, posing a heavy threat to the safety and stability of power transmission system[4]. Thus, it is urgent to investigate an effective nondestructive testing method for the internal conductive defects of composite insulators.

Fig. 1. Broken composite insulator caused by internal conductive defects

Recently, several non-destructive testing (NDT) methods are employed for detecting of inner failure of composite insulators, spanning the large spectrum, such as ultrasonic testing, X-ray, electric field measurement, infrared measurement, and ultraviolet imaging. However, these methods have their own limitation. For instance, coupling agent should be applied when using ultrasonic method[5], X-ray diffraction method process is time-consuming and harmful to human body[6], the electric field measurement method is difficult to locate the defect[7], infrared temperature measurement and ultraviolet imaging method are highly susceptible to interference testing environment[8,9]. Besides, these methods are not effective in detecting inner conductive defects in composite insulators. In recent years, some new nondestructive testing methods are being used to detect inner conductive defects of composite insulators, such as nonlinear ultrasonic[10], terahertz wave contrast method[11], optical pulse thermography[12], laser shearography[13], which brings new ideas for detecting inner defects of composite insulators.

Eddy current thermography (ECT), combining the superiorities of eddy current testing technology and infrared thermography, was proposed since the 1990 s[14]. Eddy current thermography can realize the deep internal defects in connective material such as metal fast, noninvasively and accurately. A diagram of ECT is presented in Fig. 2. A period of eddy current generated by the external high-frequency electromagnetic induction excitation unit is guided to the coil above the specimen, which will generate Joule heat inside the specimen. eddy current distribution or internal thermal conductivity will be affected for the existence of faults, which will reflect on temperature field distribution of the surface. Thermal image sequence of the temperature distribution of the surface, captured by infrared thermal imager, is analyzed by specific processing method for the purpose of the detection of defects detailly.

In the previous work, ECT has been applied to the detection of non-conductive defects (like surface cracks) of conductive materials (such as alloy)[15,16]. C. Xu et al. used ECT method for the detection of undersurface defects of GFRP composites, which applied a steel unit for the induction heat generation[17]. Y. He et al. used ECT method with excitation of pulse and lock-in modes for the purpose of the detection of photovoltaic cells and modules, achieving visual measuring of flaws in photovoltaic cells and modules[18]. Despite eddy current thermography has great effect in the detection of conductive materials, the detection of conductive defects in nonconductive materials has not been reported. Conductive defects are an important source of accidents of composite insulators. Therefore, it is urgent to investigate the feasibility of nondestructive testing of internal conductive defects by using ECT.
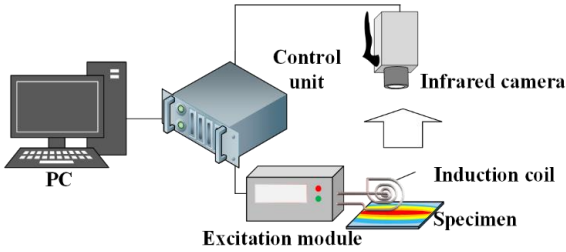

Fig. 2. Schematic of ECT System.

In present work, ECT is applied for the detection of the inner conductive flaws of composite insulators. Principal component analysis (PCA) and independent component analysis (ICA) are applied to improve the detection effect. Firstly, in section II, the principle of ECT is deduced mathematically ,and the principles of PCA and ICA are proposed. Besides, experimental step is introduced in section III. In section IV, thermal image sequence of composite insulator specimen is analyzed preliminarily. Moreover, tensor decomposition method (such as PCA, ICA) is applied to enhance the display of defects. Finally, conclusion is examined in Section V.

II. METHODOLOGY AND THEORY

A. Eddy Current Thermography

The physical process of ECT can be roughly divided into two processes, namely electromagnetic heating process and heat conduction process. In electromagnetic heating process, eddy current is generated in the material tested under electromagnetic induction, which will generate joule heat due to the resistance effect inside the material. During the heating period, joule heating power Q can be described as:

$$Q = \frac{1}{\sigma}\left|J_e\right|^2 \tau = \frac{1}{\sigma}\left|\sigma E\right|^2 \tau \tag{1}$$

Where E is defined as electric field intensity, J_e is defined as current density, σ is defined as electrical conductivity, while τ is excitation loading time. The joule heat generated by electromagnetic heating will propagate within the material, and the propagation process follows the heat conduction equation. The process above can be described as[19]:

$$\rho C_p \frac{\partial T}{\partial t} - k\nabla^2 T = Q \tag{2}$$

Where T is the temperature, k is the thermal conductivity, C_p is regarded as the specific heat, ρ is defined as the density, and t is the heat conduction time. The change of thermal diffusion will be reflected on the surface of the specimen. Besides, the basic law of infrared radiation is described as follows, which is called Stephen-Boltzmann law:

$$j = \sigma_{sb}\varepsilon T^4 \tag{3}$$

Where σ_{sb} reflects the Stephen-Boltzmann constant, ε is emissivity of the specimen, and T reflects the absolute temperature of the specimen.

B. Principle of PCA, ICA

PCA is a statistical means that converts the original data into a set of linearly independent variables, which is called the principal component[20]. The goal of this process is to retain as much primary variable information as possible by these principal components transformed from original data. Besides, major influencing factors of original variable are extracted, and the initial higher dimensional space data is transformed into lower dimensional space data, which make data easier to visualize. In effect, PCA method can be regarded as a set of linear equations, that is, to acquire $l\,(l<q)$ some variables new y_i, each variable is a linear array of the initial variable x_i, which can be described as follows:

$$\begin{cases} y_1 = m_{11}x_1 + m_{12}x_2 + \cdots + m_{1q}x_q \\ y_2 = m_{21}x_1 + m_{22}x_2 + \cdots + m_{2q}x_q \\ \qquad\qquad\vdots \\ y_l = m_{l1}x_1 + m_{l2}x_2 + \cdots + m_{lq}x_q \end{cases} \tag{4}$$

The essential purpose of ICA is for searching for the optimal linear transformation of the observed mixed signals for the purpose of the separation of independent source signals, which is based on the statistical independence of origin data[21]. Each independent component can reflect a specific data change trend. The observed signal $\mathbf{Y}(t)$ captured by the thermal imager should be standardized to generate n-dimensional vector with the zero mean $\mathbf{Y}'(t) = [y_1', y_2',..., y_n']$, which is linearly aliased from n new independent signals with zero mean $\mathbf{S}(t) = [s_1, s_2,..., s_n]$. This processing is able to be shown as follows:

$$\mathbf{Y}'(t) = \mathbf{BS}(t) = \sum_{i=1}^{n} b_i s_i \quad (i=1,2,\cdots,n) \tag{5}$$

Where $\mathbf{B}$ is a full rank matrix which is termed as mixing matrix. The ultimate goal of ICA processing is to obtain the matrix of linear transformation $\mathbf{W}_{ICA}$, which can transform independent source signal as:

$$\mathbf{X}(t) = \mathbf{W}_{ICA}\mathbf{Y'}(t) = \mathbf{W}_{ICA}\mathbf{BS}(t) \qquad (6)$$

C. Thermal Conduction from Inner Conductive Defects

In this work, inner conductive defect on the interface of composite insulator can be regarded as heat source, which is generated by induction heating, while normal nonconductive material in composite insulator won't be affected by the eddy current induction. After the heating period, thermal conduction in the composite insulator can be divided as at least two parts, which are respectively direct heating area and thermal diffusion area. As shown in Fig. 3, these two parts have different thermal distribution, which can be reflected on the surface of composite insulator. However, the difference of these two parts may be difficult to sperate, for false defect areas affected by thermal diffusion area may appear. Therefore, the separation of true and false defect areas can effectively improve the conductive defects detection effect of composite insulator.
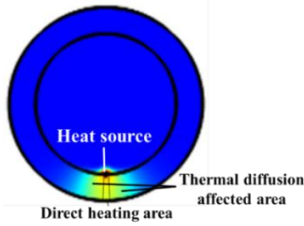

Fig. 3. Different thermal distribution of composite insulator.

III. EXPERIMENTAL SYSTEM AND SAMPLE

A. Experimental System

The experimental system can be seen in Fig. 4, which is composed of electromagnetic induction heater, water cooling system, infrared thermal camera and the control and data processing system. The electromagnetic induction heater can generate eddy currents with oscillation frequency of 100-150 kHz. Beside, the input voltage range of the electromagnetic induction heater is from 180 V to 240 V, while maximum working current of that is 1 kA. The water cooling system is provided with 0.2-0.3 MPa cooling water pressure for cooling the induction coil. The infrared thermal camera with the band range of spectral response from 1.5 m to 5.1 m, which is Telops FAST M200, carries the infrared radiation into the control and data processing system, and the resolution of the camera is able to attain 640×512 pixels. The frame frequency of the infrared thermal camera in the testing is up to 210 Hz, besides its highest frame frequency can reach 5600 Hz in window mode. The control and data processing system is integrated with industrial computer, digital silicon box, control box, display and keyboard. Data processing software is a secondary development program based on RevealIR, which can display transmitted data in real time and processing origin data to attain more accurate failure recognition results.
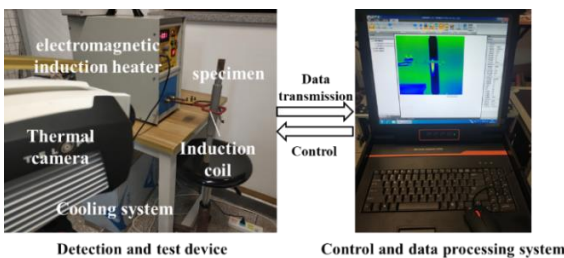

Fig. 4. The experimental system.

B. Sample

Composite insulators are mainly composed of epoxy glass fiber reinforced plastic (that is the main material of mandrel) and silicone rubber (that is the main material of sheath and umbrella skirt). During production and on-site operation, the conductive defects may exist at the sheath or the interface part between the core rod and the sheath. For the sake of simulating the conductive defects of composite insulators, steel needle is buried in the middle of the composite insulator at the interface between the core rod and sheath artificially, as shown in Fig. 5. Composite insulator sample without umbrella skirt has a glass steel core rod with the diameter of 16 mm, while the silicone rubber sheath's thickness is 5 mm. The steel needle buried in the composite insulators has a length of 80 mm and a diameter of 1.8 mm.
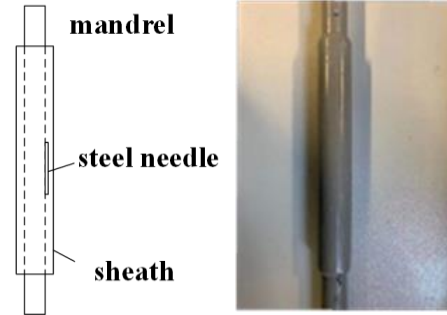
Fig. 5. The composite insulator sample.

IV. RESULTS AND DISCUSSION

A. Analysis of Original Thermal Image Sequence

In the test, the electromagnetic induction heater generated an eddy current with frequency of 100 kHz, and input current of 600 A, while the excitation loading time is 0.5 s. The infrared thermal camera captured the surface temperature of sample insulator in real time, with a frequency of 30 Hz and a duration of 40 s. Some thermal images from original thermal sequence are shown in Fig. 6. It can be concluded form Fig. 6 that high-temperature spots will appear on the surface of sample after heating phase, while high-temperature spots will gradually diffuse and the boundary of the these spots will fuzzy. This dynamic process can be seen from Fig. 6. In the thermal image of frame 120, the defect area is just showing up, while defects area is shown clearly in the thermal image of frame 170. As thermal wave conduct, the contour of defect area begins to diffuse, like the image shown in Fig. 6(c). At frame 800, the boundary of defect area is blurred, which can be seen in Fig. 6(d).

In detection of conductive material with nonconductive defects using eddy current thermography method, the detection effect would start to be better and then become worse with the passage of time[22]. However, it is different when this method is used to the detection of conductive defects of nonconductive material. Fig. 7 shows thermal responses of different point from the specimen composite insulator under testing. It can be conducted that there are at least two different thermal responses in this specimen, for point A and D are the direct heating areas, while point B and C are heat diffusion affected areas. By combining Fig. 6 and Fig. 7, the whole thermal response process can be conduct. That is, the real defect area will appear first, and the false defect area will appear as time goes on with the influence of lateral thermal diffusion. This phenomenon will blur the real defects and affect the final real detection effect.

(a) 120th frames   (b) 170th frames

(c) 360th frames   (d) 800th frames

Fig. 6.   Origin thermal images with different frames.

Fig. 7.   Thermal responses of different point.

B. PCA and ICA of ECT

As analyzed above, detection effect will become worse with the increase of testing time. Methods combined whole-process thermal image will solve this problem. In order to acquire better detection effect, PCA and ICA methods have been applied in this task.

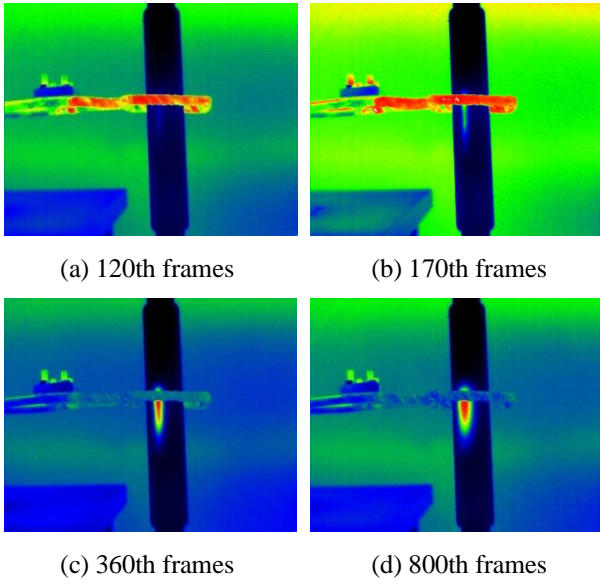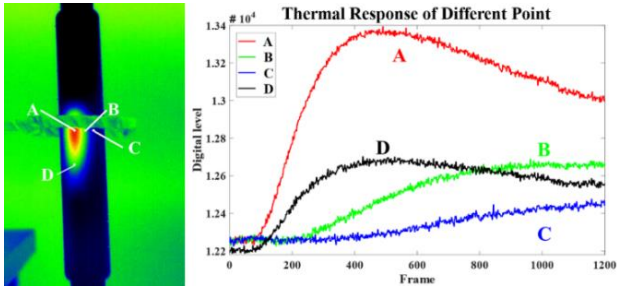In the detection test, the original thermal image sequence collected by the infrared thermal camera is composed of thermal image of P frame M×N pixels, which constitutes the three-dimensional tensor of M×N×P. When the tensor in three dimension is analyzed and processed, each frame image should be vectorized so as to transform the original three-dimensional tensor into a matrix of MN×P in two dimension. Each column of the matrix in two dimension corresponds to a frame of the image sequence, and each row corresponds to the temporal response of infrared radiation of a pixel point. Obtained two-dimensional matrix can be analyzed by PCA and ICA methods. The results of PCA are shown in Fig. 8, while the results of ICA are shown in Fig. 9.

As shown in Fig. 8, four principal components are extracted to enhance the display of different areas. In first principal component, the overall thermal image display is very similar, while the defect area and part of the heat diffusion area are highlighted. In the second principal component, the amplitude of coil is the highest, and the amplitude of defect and thermal diffusion area is the lowest. In the third principal component, the thermal diffusion area which is near the defect area has the highest amplitude, while the lowest amplitude is located at conductive defect area, thus highlighting the display of the defect. In the fourth principal component, the amplitude

of coil and background is the highest, while the amplitude of thermal diffusion affected area is lower. It can be seen that, both the first and the second principal components highlight the location of defects, but these two principal components highlight in different ways. Besides, the second principal component also highlights the display of the coil. In addition, both the second principal component and the fourth principal component highlight the thermal diffusion area. It was found that the principal component analysis did not separate the defect areas well.

(a) 1st PC   (b) 2nd PC

(c) 3rd PC   (d) 4th PC

Fig. 8.   Results of PCA.

(a) 1st IC   (b) 2nd IC

(c) 3rd IC   (d) 4th IC
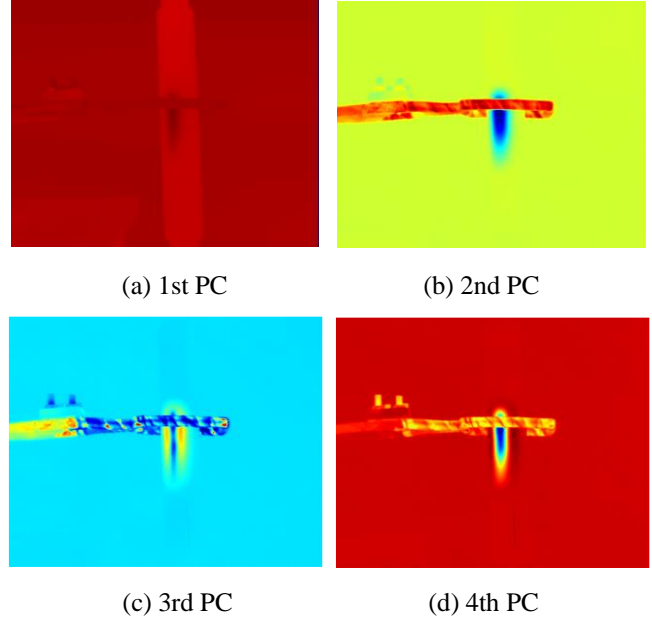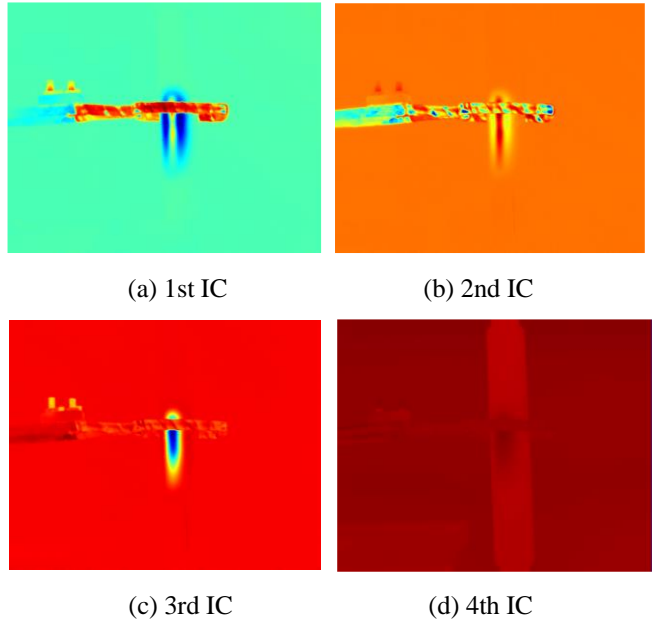
Fig. 9.   Results of ICA.

As shown in Fig. 9, different independent component shows different thermal image. In independent component 1, the amplitude of the excitation coil is higher, while the amplitude of conductive defect area and thermal diffusion area is lower. In independent component 2, the amplitude of the conductive defect field is the highest, the amplitude of the heat diffusion region is higher. Among independent component 3, the amplitude of the thermal diffusion area is the highest.

besides, the amplitude of the coil and other areas is high, while the amplitude of the conductive defect region and near defect thermal diffusion field is very low. It can be concluded from the comprehensive analysis that independent component 1 strengthened the part of the excitation coil, independent component 2 strengthened defect area, and independent component 3 strengthened peripheral thermal diffusion area.

Comparing the results of PCA and ICA, it can be found that most of the results of PCA can enhance the thermal diffusion area, but do not separate each area well. The results of ICA can separate the results of each area well, while the enhancement of defect areas is not as good as that of PCA. The reason for this phenomenon may be: (1) PCA decomposed the original data in an orthogonal way, while the true distribution of the data may not conform to the orthogonal distribution. (2) The transient response of thermal wave of the coil is mixed with the transient response of thermal wave of the insulator, which affects the data distribution processed by PCA. (3) PCA made the variance of its projection on each PC as large as possible, enhancement of the data conforms to the original statistical distribution, while ICA seeks the optimal linear variation of data, and the results obtained do not necessarily conform to the original. (4) Images are vectorized both in PCA and ICA, resulting the spatial distribution of images is ignored.

V. CONCLUSION

In this work, ECT is explored to conductive defects detection visually. Due to the influence of thermal diffusion, the detection effect of defects will gradually deteriorate with the increase of testing time. Besides, different data processing methods can improve the detection effect of defects. As for our work, the practicability of ECT for composite insulator conductive flaws is verified, through the analysis of original thermal image sequence combined with PCA and ICA methods. The main conclusions of this work are as follows: (1) ECT can be applied to identify the inner conductive flaws of composite insulators, with fast detection speed in small several decades seconds and intuitive detection results. (2) The display of internal conductive defects will be gradually blurred by the influence of thermal diffusion. (3) PCA and ICA can be used to improve the detection effect of defects, where PCA can enhance the display of defects better and ICA can separate each area better. In future research, the ECT technology could be employed in other fields related with conductive defects in nonconductive material; more efficient pattern separation methods will be explored. Besides, future research will also concentrate on inspecting different conductive materials in composite insulator for the purpose of the application for in-situ composite insulator detection.

REFERENCES

[1] Z. Guan, G. Peng, L. Wang, Z. Jia, and R. Zhang, "Application and key technical study of composite insulators," High Voltage Engineering, vol. 37, no. 3, pp. 513-519, Mar. 2011.

[2] S. M. Gubanski, A. Dernfalk, J. Andersson, and H. Hillborg, "Diagnostic methods for outdoor polymeric insulators," IEEE Trans. Dielectr. Electr. Insul., vol. 14, no. 5, pp. 1065-1080, Oct. 2007.

[3] J. Andersson, S. M. Gubanski and H. Hillborg, "Properties of interfaces between silicone rubber and epoxy", IEEE Trans. Dielectr. Electr. Insul., vol. 15, no. 5, pp. 1360-1367, Oct. 2008.

[4] X. Liang, W. Bao and Y. Gao, "Decay-like fracture mechanism of silicone rubber composite insulator," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 25, no. 1, pp. 110-119, Feb. 2018.

[5] H. Deng, Z. He and L. Chen, "Ultrasonic guided wave-based detection of composite insulator debonding," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 24, no. 6, pp. 3586-3593, Dec. 2017.

[6] Hong Yu, Jie Wei, Xianping Zhao, Yi Ma and Lei Chen, "The perspective detection of the X-ray digital radiography for the electrical equipment," 2012 International Conference on High Voltage Engineering and Application, Shanghai, 2012, pp. 591-599.

[7] H. Wang, C. Zhuang, R. Zeng, S. Xie and J. He, "Transient Voltage Measurements for Overhead Transmission Lines and Substations by Metal-Free and Contactless Integrated Electro-Optic Field Sensors," IEEE Transactions on Industrial Electronics, vol. 66, no. 1, pp. 571-579, Jan. 2019.

[8] R. Liu, Q. Shen, H. Liu, Y. Duan, Y. Zhang and C. Zhou, "Discussion on infrared accurate temperature-measuring technology of composite insulators for overhead lines," The Journal of Engineering, vol. 2019, no. 16, pp. 1759-1762, 3 2019.

[9] S. Wang, F. Lv and Y. Liu, "Estimation of discharge magnitude of composite insulator surface corona discharge based on ultraviolet imaging method," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 21, no. 4, pp. 1697-1704, August 2014.

[10] H. Wang, L. Cheng, R. Liao, S. Zhang and L. Yang, "Nonlinear ultrasonic nondestructive detection and modelling of kissing defects in high voltage composite insulators," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 27, no. 3, pp. 924-931, June 2020.

[11] Z. Zhang, L. Wang, H. Mei and W. Wu, "Quantitative Detection of Interfacial Air Gap in Insulation Equipment Based on Terahertz Wave Contrast Method," IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 12, pp. 4896-4905, Dec. 2019.

[12] L. Liu, H. Mei, C. Guo, Y. Tu, L. Wang and J. Liu, "Remote Optical Thermography Detection Method and System for Silicone Polymer Insulating Materials Used in Power Industry," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 8, pp. 5782-5790, Aug. 2020.

[13] L. Liu, C. Guo, L. Wang and H. Mei, "Nondestructive Visualization and Quantitative Characterization of Defects in Silicone Polymer Insulators Based on Laser Shearography," IEEE Sensors Journal, vol. 19, no. 15, pp. 6508-6516, 1 Aug.1, 2019.

[14] Yunze He, Gui Yun Tian, Mengchun Pan, and Dixiang Chen, "Eddy current pulsed phase thermography and feature extraction," Appl. Phys. Lett., vol. 103, no. 8, pp. 084104, Aug. 2013.

[15] Q. Yi, H. Malekmohammadi, G. Y. Tian, S. Laureti and M. Ricci, "Quantitative Evaluation of Crack Depths on Thin Aluminum Plate Using Eddy Current Pulse-Compression Thermography," IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 3963-3973, June 2020.

[16] J. Wu, J. Zhu, H. Xia, C. Liu, X. Huang and G. Y. Tian, "DC-Biased Magnetization Based Eddy Current Thermography for Subsurface Defect Detection," IEEE Transactions on Industrial Informatics, vol. 15, no. 12, pp. 6252-6259, Dec. 2019.

[17] C. Xu, W. Zhang, C. Wu, J. Xie, X. Yin and G. Chen, "An improved method of eddy current pulsed thermography to detect subsurface defects in glass fiber reinforced polymer composites," Composite Structures, vol. 242, pp. 1-14, June. 2020.

[18] Y. He, B. Du and S. Huang, "Noncontact Electromagnetic Induction Excited Infrared Thermography for Photovoltaic Cells and Modules Inspection," IEEE Transactions on Industrial Informatics, vol. 14, no. 12, pp. 5585-5593, Dec. 2018.

[19] X. Maldague, Theory and Practice of Infrared Technology for Nondestructive Testing. New York, USA: Wiley, 2001, pp. 52-69.

[20] P. Baranowski, W. Mazurek, J. Wozniak and U. Majewska, "Detection of early bruises in apples using hyperspectral data and thermal imaging," J. Food Eng., vol. 110, no. 3, pp. 345-355, June. 2012.

[21] Aapo. Hyvärinen Juha Karhunen, Erkki Oja, Independent component analysis,1st ed. New York:J. Wiley, c2001.

[22] R. Yang, Y. He, A. Mandelis, N. Wang, X. Wu and S. Huang, "Induction Infrared Thermography and Thermal-Wave-Radar Analysis for Imaging Inspection and Diagnosis of Blade Composites," IEEE Transactions on Industrial Informatics, vol. 14, no. 12, pp. 5637-5647, Dec. 2018.

Model-based Parameter Estimation Method for Terahertz Signals

Yafei Xu
Science and Technology Park
Nanshan District
Shenzhen Academy
Xi'an Jiaotong University
Shenzhen,China
State Key Laboratory for
Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
xyf2492229210@stu.xjtu.edu.cn

Liuyang Zhang*
Science and Technology Park
Nanshan District
Shenzhen Academy
Xi'an Jiaotong University
Shenzhen,China
State Key Laboratory for
Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
liuyangzhang@xjtu.edu.cn

Xuefeng Chen
State Key Laboratory for Manufacturing
Systems Engineering
Xi'an Jiaotong University
Xi'an, China
chenxf@xjtu.edu.cn

Zhen Zhang
Nanshan District
Science and Technology Park
Shenzhen Academy
Xi'an Jiaotong University
Shenzhen,China
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
zz3119101101@stu.xjtu.edu.cn

Zhonglei Shen
Nanshan District
Science and Technology Park
Shenzhen Academy
Xi'an Jiaotong University
Shenzhen,China
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
zlshen@stu.xjtu.edu.cn

Abstract—**Parameter estimation hasn't been investigated in terahertz non-destructive testing (NDT) field due to the lack of the prior distribution knowledge of terahertz signals. In this study, a statistical model of terahertz signal is proposed to model the terahertz echo signal. The critical issue mainly focuses on how to obtain the optimal parameter estimation from the complex terahertz echo signal. Therefore, we provide two estimators: the maximum likelihood estimation (MLE) based on expectation maximization algorithm (EM), and the lasso estimation based on the sparse representation. Simulation and experiments have been implemented to analyze the fitting performance of the proposed model. The results validate the effectiveness and applicability of the statistical model, and indicate that the Lasso estimator outperforms the EM estimator for the parameter estimation of terahertz signal, which provides a new statistical distribution model and parameter estimation method for the terahertz signal in terahertz NDT.**

Keywords—*parameter estimation, terahertz NDT, prior statistical model, EM, sparse representation*

I. INTRODUCTION

Terahertz technique, as a relatively novel and promising non-destructive testing approach, has paid great attention and extensive applications in various non-polar materials with the single or multilayer structures due to its superior properties such as high time and spatial resolution, noninvasive, noncontact, and nonionizing [1]. Compared with the conventional NDT methods such as ultrasonic detection, terahertz inspection method can provide higher spatial resolution without a liquid couplant due to its short wavelength [2]. At present, there have been many reports on the applications of terahertz in NDT. For instance, Ryu et al. [3] utilized the terahertz time-domain spectroscopy to distinguish the preset hidden multi-delamination defects in a glass-fiber-reinforced plastic (GFRP) composite laminates, and validated the effectiveness of method by deducing the Fresnel equations. Fukuchi et al. [4] measured the topcoat thickness of thermal barrier coating (TBC) applied to a gas turbine blade by using terahertz waves, and the result indicated that the obtained thickness agreed with microscope observation results. Stoik et al. [5] fulfilled the localization and imaging of the voids and cracks inside aircraft composites by using terahertz detection and imaging. In addition, Dai et al. [6] improved the time resolution of terahertz non-destructive detection of debonds located in layered structures based on wavelet transform. Chen et al. [7] proposed a novel frequency-wavelet domain deconvolution (FWDD) method for the terahertz reflection imaging, which promotes the development of terahertz signal process.

Despite advances that have been made above in terahertz non-destructive testing field, there has never been an appropriate statistical model to fit the terahertz signal theoretically, and the corresponding parameter estimation method has not studied as well, which limits its further wide applications.

In the work, we propose a statistical model to fit the terahertz signal based on the Gaussian mixture model. In the process, two parameter estimation methods are provided to obtain the optimal parameter estimation of terahertz echo signal. Then, a series of simulation and experiments are implemented to analyze the fitting performance of both methods. The result indicates the effectiveness of the statistical model and provides a novel insight for the terahertz NDT.

II. MODEL AND PARAMETER ESTIMATION METHODS

A. Model of Terahertz Signals

*Liuyang Zhang is the corresponding author. (e-mail: liuyangzhang@xjtu.edu.cn).

In the pulse-echo terahertz testing, the reflected terahertz echo signal from a flat metal plate can be modeled by the Gaussian mixture model as

$$y(\theta;t) = a_1 * \exp(-(t-b_1)^2 / c_1) + a_2 * \exp(-(t-b_2)^2 / c_2) \quad (1)$$

$$\theta = [a_1 \ \ b_1 \ \ c_1 \ \ a_2 \ \ b_2 \ \ c_2]$$

where, a_1 and a_2 represent the amplitudes corresponding to a narrow (subscript 1) and a broad (subscript 2) Gaussian echo model, respectively, which indicates the weight of the mixture in the terahertz signal. b_1 and b_2 represent the initial delay time relating to the location of the reflector. c_1 and c_2 represent bandwidth factors that determine the time duration of the echo in time domain, which is associated with the attenuation and dispersion of terahertz wave in the propagation path. The parameters have intuitive physical meanings for the terahertz NDT process. Therefore, the prior distributions of the echoes for complicated sample tests can be obtained by superimposing the Gaussian mixture models via varying the parameters.
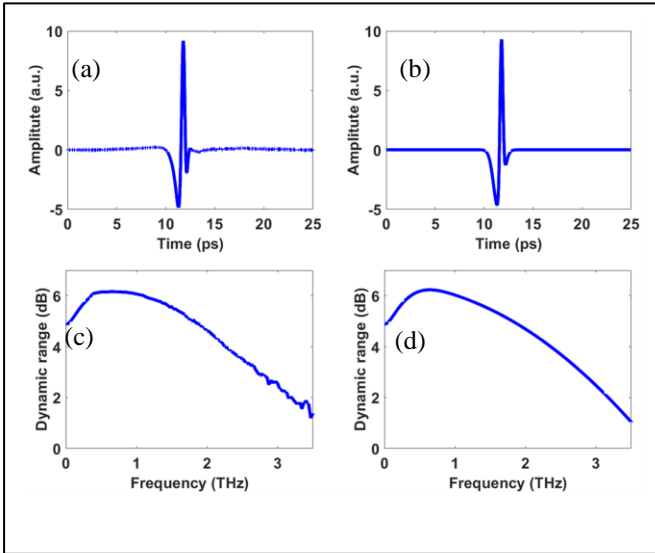

Fig. 1. (a) The measured terahertz reflected echo signal from the front surface of the metal plate; (b) The estimated terahertz echo signal; (c) The frequency spectrum corresponding to (a); (d) The frequency spectrum corresponding to (b).

To demonstrate the effectiveness and applicability of the mixture model, a parameter estimation based on the experimental data is implemented. As shown in Fig.1, the terahertz reflected echo (reference echo) in time domain from the front surface of the metal plate is illustrated in Fig.1 (a), and the corresponding frequency spectrum is shown in Fig.1 (c). Fig.1 (b) depicts the estimated terahertz echo signal in time domain from the obtained experimental data by using the a Gauss Newton (GN) algorithm. The estimated parameter θ is $[13.35 \ 11.76 \ 0.2 \ -5.039 \ 11.45 \ 0.7]$. The corresponding frequency spectrum of the estimated terahertz echo is shown in Fig.1 (d). It is concluded that the Gaussian mixture model with specific parameters can fit the terahertz echo signal well in both time domain and frequency domain.

B. *Expectation Maximization (EM) Method*

In practical terahertz NDT, the reflected echoes from the reflectors are not usually single terahertz echo, but the superimposed signal with multiple terahertz echoes. The maximum likelihood estimation based on the expectation maximization algorithm (EM) is a widely used multi-parameters estimation method in NDT field [8]. Here, we firstly introduce the method to parameter estimation of terahertz echoes. Considering the actual terahertz echo is composed of M-superimposed dispersive echoes with the white Gaussian noise (WGN), and its formulation can be expressed as

$$Y(\theta;t) = \sum_{m=1}^{M} y(\theta_m;t) + e(t) \quad (2)$$

where, $Y(\theta;t)$ is the observed signal with noise, $y(\theta_m;t)$ is the noiseless single echo component, and the $e(t)$ is the white Gaussian noise.

Based on the Gaussian mixture model, we define the Y_m as the "unobserved data" set of the mth terahertz echo:

$$Y_m = y(\theta_m;t) + e(t) \quad (3)$$

The observed signal can be obtained by the linear transformation of the unobserved data:

$$Y(\theta;t) = \sum_{m=1}^{M} Y_m \quad (4)$$

It has been proven that the MLE of the parameter vectors θ_m subjected to the unobserved data sets Y_m can be computed and the MLE of θ_m maximizes the probability density function (pdf) associated with the unobserved data set Y_m. Thus, the expectation of Y_m can be computed in terms of the observed data and the current value of the parameter vectors can be expressed as

$$\hat{Y}_m^{(k)} = y(\theta_m^{(k)}) + \gamma_m \left[Y - \sum_{i=1}^{M} y(\theta_i^{(k)}) \right] \quad (5)$$

where, $\sum_{m=1}^{M} \gamma_m = 1$, and (k) is the iteration number. This is the E-step in EM algorithm, and the maximization step (M-step) utilizes the expectation from E-step to maximize the pdf. In the M-step, the parameter vectors are updated by minimizing

$$\theta_m^{(k+1)} = \arg\min_{\theta_m} \left\| \hat{Y}_m^{(k)} - y(\theta_m) \right\|^2 \quad (6)$$

In conclusion, the EM algorithm for parameter estimation of M-superimposed terahertz echoes with WGN can be performed as the following steps: [9]

Step 1. Make initial guesses for the parameter vectors and form matrix as $\theta^{(0)} = [\theta_1^{(0)}, \theta_2^{(0)}, \theta_3^{(0)}, ..., \theta_M^{(0)},]$, then set the iteration number $k = 0$ and $k_{max} = A$;

Step 2. (E-step) For $m = 1, 2, 3, ..., M$, compute the (5), and obtain the expected terahertz echoes. Here, $\gamma_m = 1/M$;

Step 3. (M-step) For $m = 1, 2, 3, ..., M$, iterate the (6) and set $\theta_m^{(k)} = \theta_m^{(k+1)}$;

Step 4. Check the convergence criteria and iteration number maximum, if $\left\| \theta_m^{(k+1)} - \theta_m^{(k)} \right\| \leq$ tolerance or $k \geq A$, then stop;

Step 5. Set $k \to k+1$ and go to Step 2.

C. Lasso-based Parameter Estimation Method

Another multi-parameters estimation method proposed for terahertz superimposed echoes is the least absolute shrinkage and selection operator (Lasso) based on the sparse representation. Since the measured THz reflected echo signal $Y(t)$ can be obtained by the convolution between the THz reference signal $y_{ref}(t)$ and the impulse response function $h(t)$ with extra Gaussian noise $e(t)$, the convolution model can be expressed as

$$Y(t) = y(t) \otimes h(t) + e(t) \tag{7}$$

In practice, we should consider the discrete form of (7) with the sampling period T

$$Y_n = \sum_{k=0}^{N-1} y_{n-k} h_k + e_n \tag{8}$$

where $Y_n = Y(nT)$, $h_k = h(kT)$. n and k represent the indices of data sampling points, and N represents the length of measured signal.

Considering the sparse prior of the impulse response function, the sparse decomposition model of measured terahertz echo can be established as

$$Y = \phi h + e \tag{9}$$

where, ϕ represents the overcomplete dictionary composed of the parameter vectors $\theta = [\tau, c_1, c_2]$, τ represents the delay time, c_1 and c_2 represent the bandwidth factor. h is the sparse vector to be recovered, i.e., it has only few nonzero components, which corresponds to the location of the parameters in the dictionary ϕ. Therefore, in terahertz signal process, the parameter estimation problem can be transformed into the recovery problem of the impulse response function from the complex measured terahertz echoes, and the kernel issue is how to recover the impulse response sequence from the superimposed echoes with sparsity and fidelity.

In this case, the sparse vector h can be calculated accurately by solving the l_0 regularized optimization problem, which is expressed as

$$\hat{h} = \arg\min_h \| h \|_0 \quad \text{s.t.} \quad \| Y - \phi h \|_2 \leq \sigma \tag{10}$$

where the l_0 - norm of $\| h \|_0$ denotes the number of nonzero elements in the vector h. σ is the regularization parameter with respect to the noise level.

It is known that the global optimal cannot be guaranteed when l_0 regularized optimization problem is solved as shown in (10) due to its nonconvexity [10]. It has been demonstrated that the nonconvex optimization problem of l_0 - norm can be addressed by the convex optimization method such as Lasso

[11]. Thus, equation (10) can be described as Lasso problem as

$$\hat{h} = \arg\min_h \| Y\text{-}\phi h \|_2 \quad \text{s.t.} \quad \| h \|_1 \leq \varepsilon \tag{11}$$

where $\|\cdot\|_2$ represents the l_2 - norm. $\|h\|_1$ represents the l_1 - norm of h. ε represents the regularization parameter that balances the sparsity and fidelity.

Therefore, the sparse impulse response sequence can be obtained by solving (11) using a spectral projected-gradient (SPG) algorithm in ref 12]. The parameter vectors of measured signal can be also obtained from the obtained h.

III. SIMULATION

In this section, a series simulation are performed to analyze the performance of EM and Lasso for the parameter estimation of superimposed echo signal in terahertz NDT. It should be noticed that the present work can be easily extended to more complicated cases.

Initially, we form a synthetic terahertz reflected signal composed of three Gaussian mixture models with the dispersion and attenuation to simulate the actual measured signal as much as possible. The initial parameter vectors θ of synthetic echo signal are set to $\theta = [\theta_1, \theta_2, \theta_3]$, detailed parameters as shown in TABLE I. The synthetic terahertz echo signal contains 4096 data points with a sampling frequency $F_s = 128$, as shown in Fig. 2.

TABLE I. INITIAL TRUE PARMETER VECTORS

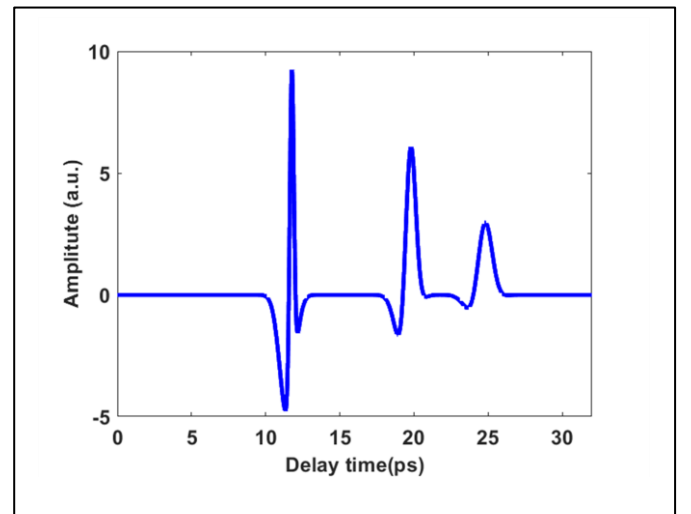| True parameter	a_1	b_1	c_1	a_2	b_2	c_2
θ_1	13.35	11.76	0.2	-5.039	11.45	0.7
θ_2	8.9	19.76	0.5	-3.359	19.45	0.8
θ_3	4.45	24.76	0.7	-1.68	24.45	1

Fig. 2. The synthetic terahertz echo signal.

Based on the model (1) of terahertz signal, both EM and Lasso method are employed to estimate the optimal model

parameter. For EM method, an initial guess for the parameter vectors is necessary, and the estimation accuracy depends on the degree of the initial guess close to the true value. Here, we set the initial guess parameter vector as TABLE II.

TABLE II.   INITIAL GUESS PARMETER VECTORS

Initial parameter	a_1	b_1	c_1	a_2	b_2	c_2
θ_1	12	11.5	0.15	-5	11	0.7
θ_2	9	10	0.5	-3	10	0.7
θ_3	4.5	22	0.6	-2	22	0.9

As shown in TABLE III, the results obtained by EM indicate that the estimation is close to unbiased estimation when the measured signal is noise-free. However, when the signal to noise ratio (SNR) decreases gradually, the estimation accuracy for the amplitude parameter $[a_1, a_2]$ and the bandwidth factor $[c_1, c_2]$ is the same as expected, whereas the estimation accuracy for the delay time $[b_1, b_2]$ deteriorates, that is to say, the estimation performance degrades gradually with the increment of the noise level due to the distortion of the actual echo parameter caused by the noise, which indicates the sensitivity of EM to the added noise.

For Lasso method based on the sparse representation, an initial guess for the parameter vectors is not necessary, and its estimation accuracy depends on the construction of a dictionary and the selection of the regularization parameter. Therefore, we design the overcomplete dictionary based on the initial guess range of the parameter vectors. Considering the sparse prior knowledge of the impulse response sequence, the estimated parameter vectors are illustrated in TABLE IV.

In TABLE IV, it can be observed that the overall estimation performance of the parameter vectors is good, especially for the delay time parameters and the bandwidth factors. In addition, the method is not affected by the added noise, which indicates the potential of Lasso parameter estimation method for the time of arrival (TOA) of echo and the location of the defects (LOD).

To quantitatively analyze the estimation performance of both methods, the root-mean-square error (RMSE) is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\hat{Y}(j)-Y(j))^2} \qquad (12)$$

where, N is the length of terahertz echo signal. $\hat{Y}$ and Y represent the reconstructed echo signal and the origin echo signal, respectively. j represents the index of the data points.

In the analysis process, we utilize a Monte-Carlo simulation to compute the RMSE of both method with the same noise level in 1000 trials. The results are illustrated in the right columns of TABLE III and TABLE IV, which reveals the Lasso method outperforms the EM method for the parameter estimation of terahertz signal with different noise levels.

TABLE III.   THE ESTIMATED PARAMETER VECTORS OBTAINED BY EM

SNR (dB)	a_1	b_1	c_1	a_2	b_2	c_2	RMSE
No-noise	13.35	11.76	0.2	-5.039	11.45	0.7	0
	8.9	16.76	0.5	-3.359	16.45	0.8	
	4.45	21.76	0.7	-1.68	21.45	1	
20dB	13.3	11.76	0.2	-5.03	11.5	0.7	0.0021
	8.88	16.76	0.5	-3.34	16.5	0.8	
	4.16	21.77	0.69	-1.41	21.4	1.04	
10dB	13.48	11.76	0.2	-5.13	11.46	0.71	0.0066
	8.2	16.77	0.48	-2.72	16.38	0.81	
	4.39	21.76	0.67	-1.52	21.4	0.98	

TABLE IV.   THE ESTIMATED PARAMETER VECTORS OBTAINED BY LASSO

SNR (dB)	a_1	b_1	c_1	a_2	b_2	c_2	RMSE
No-noise	13.34	11.76	0.2	-5.036	11.45	0.7	0
	8.86	19.76	0.5	-3.34	19.45	0.8	
	4.44	24.76	0.7	-1.67	24.45	1	
20dB	13.31	11.76	0.2	-5.03	11.45	0.7	0.002
	8.72	19.76	0.5	-3.29	19.45	0.8	
	4.39	24.77	0.7	-1.66	24.45	1	
10dB	12.89	11.76	0.2	-4.87	11.45	0.71	0.0065
	8.19	19.76	0.5	-3.09	19.45	0.81	
	3.74	24.76	0.7	-1.41	24.45	0.98	

IV. EXPERIMENTS

To further demonstrate the effectiveness of the proposed model and the performance of both estimation methods, a three-layer thin sample that consists plastic layer, adhesive layer and metal layer is measured by the reflective terahertz time-domain spectra system under constant temperature and humidity environment. The reflected terahertz echo signal of the sample is obtained with the sampling period $T=0.0081ps$, sampling data points $N=4096$, as shown in Fig.3 (a). It can be found that the measured echo signal is a complex superposition signal that is composed of the effective echo signal, multiple interlayer reflection signals and additive noise signal. The effective echo signal in measured signal can be reconstructed by EM and Lasso methods based on the Gaussian mixture model. It is worth noting that the multiple interlayer reflection signals that are meaningless to signal analysis are ignored during the parameter estimation process.

TABLE V is the initial guess values of the parameter vectors. TABLE VI shows the estimated parameter vectors from the EM method and Lasso method. It can be found that both method is almost identical in estimating the delay time of first two wave packets, however, for the third wave packet, there are a significant difference, as shown in Fig.3 (b) and (c), which can be attributed to the selection of the initial value for EM method.

In order to quantitatively analyze the fitting performance of the proposed model, RMSE and fitting coefficient (R^2) is employed.

$$R^2 = 1 - \frac{S_{res}}{S_{tot}}$$

$$S_{res} = \sum_i \left(Y_i - \hat{Y}_i \right)^2 \qquad (13)$$

$$S_{tot} = \sum_i \left(Y_i - \bar{Y} \right)^2$$

where, Y represents the measured echo signal. $\hat{Y}$ represents the reconstructed echo signal. $\bar{Y}$ represents the mean value of the measured echo signal. i represents the index of the data points.

As illustrated in Fig.3, the fitting coefficient ($R^2 = 0.9679$) for Lasso method is higher than that ($R^2 = 0.9401$) for EM method, and RMSE (0.1340) for Lasso method is lower than that (0.1829) for EM method, which reveals that the effectiveness of the proposed terahertz theoretical model, and the superiority of the corresponding Lasso method for the parameter estimation of terahertz signal.

TABLE V.   INITIAL GUESS PARMETER VECTORS FOR EM METHOD

Initial parameter	a_1	b_1	c_1	a_2	b_2	c_2
θ_1	7.6	13.45	0.23	-3	13.12	0.7
θ_2	6.5	15.75	0.35	-3	15.42	0.9
θ_3	1	18.45	0.8	-2	18.12	0.5

TABLE VI.   THE COMPARISION OF ESTIMATED PARAMETER VECTORS OBTAINED BY EM AND LASSO METHOD

Method	a_1	b_1	c_1	a_2	b_2	c_2	RMSE
EM	7.64	13.45	0.22	-2.91	13.12	0.7	0.1829
	5.81	15.78	0.33	-3.23	15.45	0.74	
	0.26	18.52	4.71	0.83	18.19	0.30	
Lasso	7.76	13.44	0.23	-2.99	13.13	0.67	0.1340
	6.54	15.79	0.36	-3.74	15.55	0.78	
	-1.46	18.11	0.34	1.07	17.26	0.82	

V. CONCLUSION

In this work, we first propose a statistical model of terahertz echo signal based on the prior knowledge of the measured terahertz signal. and the effectiveness in time domain and frequency domain. To demonstrate the effectiveness of the proposed model, both optimal parameter estimation method such as the expectation maximization method and the lasso-based method are employed. Then, numerical simulation and experiment are conducted to analyze the fitting performance of the model by both methods for the optimal parameters estimation of terahertz echo signal.

The results validate the effectiveness of the proposed model and indicate that the Lasso estimator outperforms the EM estimator for the parameter estimation of terahertz signal, which provides a new statistical distribution model and parameter estimation method for the terahertz signal in terahertz NDT.
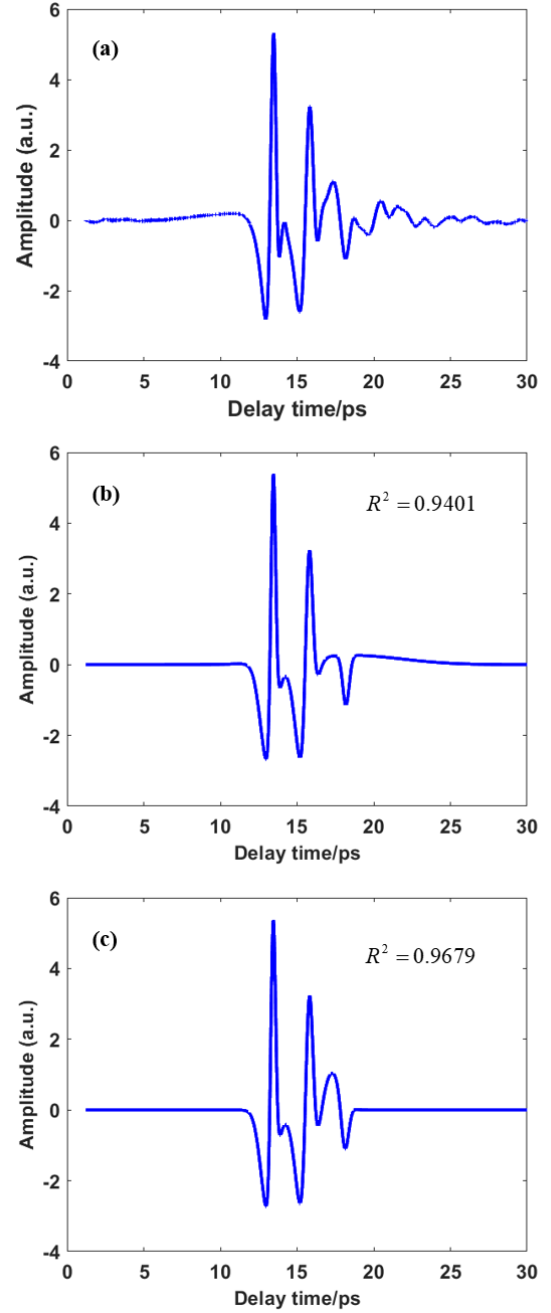

Fig. 3. The comparison of the fitting performance between the EM and Lasso methods. (a) The measured terahertz echo signal; (b) The reconstructed echo signal by EM; (c) The reconstructed echo signal by Lasso.

REFERENCES

[1] A. Redo-Sanchez et al., "Terahertz time-gated spectral imaging for content extraction through layered structures," Nature Communications, vol. 7, no. 1, p. 12665, 2016/09/09 2016.

[2] C. Jördens, S. Wietzke, M. Scheller, and M. Koch, "Investigation of the water absorption in polyamide and wood plastic composite by terahertz time-domain spectroscopy," Polymer Testing, vol. 29, no. 2, pp. 209-215, 2010/04/01/ 2010.

[3] C.-H. Ryu, S.-H. Park, D.-H. Kim, K.-Y. Jhang, and H.-S. Kim, "Nondestructive evaluation of hidden multi-delamination in a glass-fiber-reinforced plastic composite using terahertz spectroscopy," Composite Structures, vol. 156, pp. 338-347, 2016/11/15/ 2016.

[4] T. Fukuchi et al., "Topcoat Thickness Measurement of Thermal Barrier Coating of Gas Turbine Blade Using Terahertz Wave," Electrical Engineering in Japan, vol. 189, no. 1, pp. 1-8, 2014/10/01 2014.

[5] C. Stoik, M. Bohn, and J. Blackshire, "Nondestructive evaluation of aircraft composites using reflective terahertz time domain spectroscopy," NDT & E International, vol. 43, no. 2, pp. 106-115, 2010/03/01/ 2010.

[6] B. Dai et al., "Improved terahertz nondestructive detection of debonds locating in layered structures based on wavelet transform," Composite Structures, vol. 168, pp. 562-568, 2017/05/15/ 2017.

[7] Y. Chen, S. Huang, and E. Pickwell-MacPherson, "Frequency-Wavelet Domain Deconvolution for terahertz reflection imaging and spectroscopy," Optics express, vol. 18, pp. 1177-90, 01/18 2010.

[8] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 36, no. 4, pp. 477-489, 1988.

[9] H. Jia, Z. Zhang, H. Liu, F. Dai, Y. Liu, and J. Leng, "An approach based on expectation-maximization algorithm for parameter estimation of Lamb wave signals," Mechanical Systems and Signal Processing, vol. 120, pp. 341-355, 2019/04/01/ 2019.

[10] H. Zhang, X. Chen, Z. Du, and R. Yan, "Kurtosis based weighted sparse model with convex optimization technique for bearing fault diagnosis," Mechanical Systems and Signal Processing, vol. 80, pp. 349-376, 2016/12/01/ 2016.

[11] E. van den Berg and M. P. Friedlander, "Probing the Pareto Frontier for Basis Pursuit Solutions," SIAM Journal on Scientific Computing, vol. 31, no. 2, pp. 890-912, 2009/01/01 2008.

[12] E. G. Birgin, J. M. Martínez, and M. Raydan, "Nonmonotone Spectral Projected Gradient Methods on Convex Sets," SIAM Journal on Optimization, vol. 10, no. 4, pp. 1196-1211, 2000/01/01 2000.

Battery State of Health Estimation with Incremental Capacity Analysis Technique

Yuanyuan Li
The School of automation engineering
University of Electronic Science and
Technology of China
Chengdu, China
liyuanyuanfy@163.com

Hanmin Sheng
The School of automation engineering
University of Electronic Science and
Technology of China
Chengdu, China
hmsheng8911@hotmail.com

Yuhua Cheng*
The School of automation engineering
University of Electronic Science and
Technology of China
Chengdu, China
yhcheng@uestc.edu.cn

Abstract—State of health estimation of lithium-ion battery based on data-driven methods are influenced by the model input. Moreover, utilizing multi-dimensional health indicators as model input to estimate battery state of health does not increase the estimated accuracy, while it will increase the computational burden of the battery model. Therefore, this paper presents a prognostic framework based on principal component analysis technique to decrease the number of model input, while the health indicators are extracted from battery incremental capacity curves. Besides, the relationship between health indicators and state of health is established by Gaussian process regression. In addition, considering the diversity of the battery operating condition, this paper analyses the prediction of battery state of health under the three aging temperature mode, which consider four verified types. Moreover, using eight lithium-ion cells which aged at three different conditions to test the performance of our approach. The proposed methods have a higher accuracy based on the results of three types of errors, which show it can get less than 2.5% estimated error.

Keywords—lithium-ion battery, state of health, principal component analysis, incremental capacity analysis

I. INTRODUCTION

Lithium-ion battery has applied in many energy storage system, such as electric vehicles, e-mobility in general, and other battery energy storage systems. Nevertheless, the lithium-ion battery performance changes with its continuous use, which will directly lead to the degradation for battery capacity and further affect the normal use for energy storage equipment. Battery management system (BMS) plays a huge role in maintaining safe reliable and efficient operation of lithium-ion battery. As an important task in BMS, it is important to monitor battery capacity degradation and optimizing charging and discharging strategy so as to obtain accurate battery state of health (SOH) estimation [1,2].

To address this problem, there has been a large amount methods of battery SOH estimation in recent years. Generally speaking, these approaches can be classified in two categories, the model-based and the data-driven methods.

The model-based method is aim at building mathematical models of battery degradation phenomenon for considering the electrochemical mechanism or the physicochemical aging mechanism. The two common models based on this method are the electrochemical model and the equivalent circuit model [3,4]. With the model-based method, the battery degradation mechanism can be clearly explained. However, the establishment of the battery model not only depends on the experimental equipment and load conditions, but also this model has high complexity and complex calculations.

Different from model-based methods, data-driven methods can ignore the internal evolution mechanism of battery capacity degradation. Based on battery data information, the battery capacity degradation model is established so as to completing the battery SOH prediction. Existing approaches include kernel ridge regression (KRR), Monte carlo, Long short-term memory networks (LSTM), and so on. Maitane *et al.* [5] have presented the battery SOH estimation prognostic framework with a low computational effort. Bayesian networks modeling battery fade with impedance has been proposed in [6].

Besides, most data-driven methods are focus on extracting the "good features" from the battery raw data, which finds the bridge between these features and battery SOH, to building the battery capacity degradation，model. As one of the most commonly feature methods, incremental capacity analysis (ICA) technique has received widespread attention [7-13]. However, there are several issues of these methods, which want to overcome: (1) Most references based on ICA technical is aim at extracting partial health indicators, as in [7,12,13]. It has not enough detailed description on the health indicators based on ICA, which lacks of explanation of incremental capacity (IC) curve characteristics. (2) In [7,8,12,13,14], they have aimed at extracting health indicators, but the redundancy of the health feature information has not been considered, it has doubted about the idea that the more features the better the prediction effect. (3) The influence of temperature on battery health cannot be ignored in real application. In [7,12,13,14], they have just extracted battery health indicators under one temperature mode, which has been shored of analysis of health indicators at different temperatures modes, the trend of health indicators is doubtful under different temperature mode.

Based on the above discussion, this paper proposed an improved GPR model for accurate battery SOH estimation. Multi-dimension health indicators are extracted from IC curves, which are extracted all health indictors based on ICA technical, as for the base choice for battery model input. However, multi-dimension health indicators will increase the complexity of the model, and does not improve the prediction accuracy. Multi-dimension health indicators are processed by the principal component analysis (PCA) algorithm, and the processed "health indicators" are used as the new model input variable to complete the battery capacity degradation model. Besides, considering the diversity of operating condition, three temperature condition will be considered in this paper. The estimated results show the high efficiency and robustness based on eight cells. The main contributions of this paper are summarized as follows:

1) Considering the different operation condition, such as charging rate, aging state and temperature, multi- health

---

*Yuhua Cheng is the corresponding author. (e-mail: yhcheng@uestc.edu.cn).

indicators based on ICA technique has been detailed in this paper.

2) In order to eliminate the negative effect of computational burden caused by multi-dimensional input, the PCA algorithm reduces the dimensionality of the model input. The SOH is estimated effectively with considering the three types errors based on our proposed method, which means the multiple health indicators do not increase the accuracy of battery SOH estimation.

3) In this paper, the training and testing date set are from different data set, which means that in the absence of cell data that needs to be predicted, other cell data is trained the model to achieve the battery SOH prediction. Based on four types compared methods, the SOH estimated result has a high accuracy.

The organization of this paper is arranged as follows. The corresponding laboratory battery tests are presented in Section II. The related algorithms including PCA and GPR are presented in Section III. The health feature extraction method based on incremental capacity analysis are shown in Section IV. Then, the section V discusses the SOH estimated results. Finally, the conclusion is summarized in Section VI.

II.   BATTERY EXPERIMENT

Fig. 1 shows the lithium-ion battery cell which used in this paper. The cell's basic electric parameters are summarized in Table I.
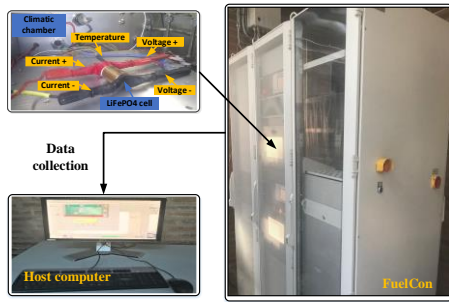

Fig. 1. The battery test bench.

TABLE I.   THE DESCRIPTION FOR THE TESTING BATTERY

Description	Value
Nominal Capacity	2.5 Ah
Charge Current	2.5A
Discharge Current	2.5A
Charge cut-off voltage	3.6V
Discharge cut-off voltage	2V

These LFP/C battery cells were placed in temperature control chambers in order to allow stable temperature during the test. The temperature values summarized in Table II, which refers to the average temperature measured on the battery cell during cycling, which not to the ambient temperature inside the various temperature chambers, which these cells were aged at these three temperature modes, 42.5℃, 35℃ and 25℃ respectively. During the reference test, the charging and discharging battery capacity was measured with 2.5 A (1C-rate) and 10 A (4C-rate) at their own temperature condition. The measured voltage and current during one reference performance test, are presented in Fig. 2. We have

eight classes battery data, which these data set can be divided into three working mode condition, the cell's working mode condition is detailed in Table II.

TABLE II.   THE WORKING MODE CONDITION FOR BATTERY DATA

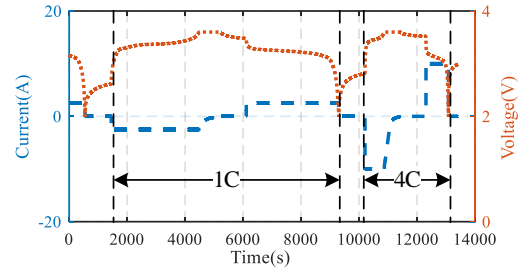| Test Case	LFP/C cells	Temperature	Working mode condition
TC1	cell 1,cell 2, cell 3	42.5℃	Mode 1
TC2	cell 4,cell 5, cell 6	35℃	Mode 2
TC3	cell 7,cell 8	25℃	Mode3

Fig. 2. The representation for the whole testing process.

In this paper, three different modes data are used to verify our proposed methods. We select three batteries in different temperature modes and observe their capacity degradation. Take cell 1, cell 4, and cell 7 as examples, the capacity evolution of the these batteries during the aging process is presented in Fig. 3.
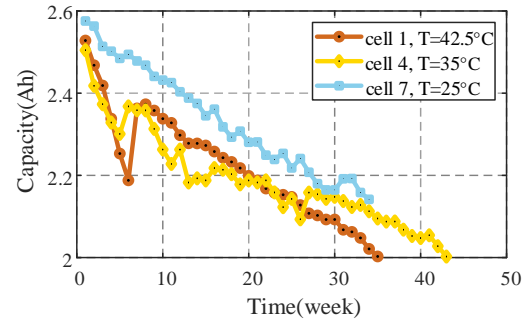

Fig. 3. The capacity variation of the three LiFePO4/C batteries.

III.   BATTERY SOH BASED ON GAUSSIAN PROCESS REGRESSION

Due to the advantages with nonparametric and probabilistic, the Gaussian process regression (GPR) model has been used in battery SOH estimation. By using the prior knowledge in Bayesian framework, while quantifying the prediction uncertainty in a principled way, the GPR model can realize to achieve the state prediction. In this paper, our goal is to establish the bridge between health indicators and battery SOH.

Consider the input vector x, output vector y and $f(x)$, in order to estimate the output y, it can be written as $f(x) \sim GP(m(x), k(x, x'))$, where $m(x) = E[f(x)]$ and $k(x, x') = E[(f(x) - m(x))(f(x') - m(x'))]$ are mean function and the covariance function respectively.

The output function $f(x)$ is unknown and normally what we obtain is the observed value of the input, as shown in (1)

$$Y = f(x) + \varepsilon \tag{1}$$

where Y is the measured value, ε is the measure noise and $\varepsilon \sim N(0, \sigma_n^2)$. Hence, the prior distribution of observations can be expressed as follows:

$$Y \sim N(m(x), k(x, x') + \sigma_n^2 I_n) \tag{2}$$

In addition, if we have original input data $X = [x_1, ..., x_n]$, original measure data $Y = [y_1, ..., y_n]$ \$, a new data X^*, the predicted value f^* can be obtained based on the original input/measure data, which can be shown as follows,

$$\begin{bmatrix} Y \\ f^* \end{bmatrix} \sim N(0, \begin{bmatrix} K(X,X) + \sigma_n^2 I_n & K(X,X^*) \\ K(X^*,X) & K(X^*,X^*) \end{bmatrix}) \tag{3}$$

where $K(X, X)$ is a covariance matrix with original input, $K(X^*, X^*)$ is a covariance matrix with new data, $K(X, X^*)$ is a covariance matrix between original input and new data, $\sigma_n^2 I_n$ is the noise covariance matrix.

Moreover, our goal is to get the predicted value f^*. Before completed the results, the focus point is to get the condition distribution on with X, Y and X^*, which can be shown as follows,

$$f^* \mid X, Y, X^* \sim N(\overline{f}^*, \text{cov}(f^*)) \tag{4}$$

where $\overline{f}^*$ is the mean value of the prediction result of f^*, $\text{cov}(f^*)$ is the covariance value of the prediction result of f^*, which are shown as follow,

$$\overline{f}^* = K(X^*,X)[K(X,X) + \sigma_n^2 I_n]^{-1}Y,$$
$$\text{cov}(f^*) = K(X^*,X) - K(X^*,X^*)[K(X,X) + \sigma_n^2 I_n]^{-1}K(X^*,X) \tag{5}$$

Besides, by using the marginal likelihood to identify the parameters, while the marginal log-likelihood can be reached the maximum value, which can be written as (6),

$$\log(Y \mid X, \Theta) = -\frac{1}{2}Y^T K_y^{-1} - \frac{1}{2}\log |K_y| - \frac{n}{2}\log 2\pi \tag{6}$$

where $K_y = K + \sigma_n^2 I_n$ refers to the co-variance matrix of Y, the Θ is the hyper-parameter set, then the maximum likelihood method can be used to determine parameters set Θ. The gradient descent algorithm is used to solved (6), which can be shown in (7)

$$\begin{cases} \dfrac{\partial}{\partial \Theta_i} \log p(Y \mid X, \Theta) = \dfrac{1}{2}tr((\alpha\alpha^{-1} - K_y^{-1}\dfrac{\partial K}{\partial \Theta_i})\dfrac{\partial K}{\partial \Theta_i}) \\ \alpha = K_y^{-1}Y = (K + \sigma_n^2 I_n)^{-1}Y \end{cases} \tag{7}$$

Based on above analysis, the battery SOH estimation with GPR model can be obtained.

IV. FEATURE PARAMETERS

Based on the measured laboratory results, in this part, we propose battery cell characteristic parameters that can track the battery SOH, which are extracted from battery experimental data with ICA technique.

The ICA technique can be used to process battery data, while the flat voltage curves are converted into a sequence of peaks and valleys, that is the incremental capacity (IC) curve [7-11]. Each IC peak and valleys of the IC curves has their own shapes, amplitudes and positions under different aging states, as presented in [8]. Thus, it is important to extracted the useful information from these external characteristics. The representation of IC curves for the tested battery at different aging states is plotted in Fig. 4. The specific calculation is shown in (8).

$$\begin{cases} \dfrac{dQ}{dV} = \dfrac{\Delta Q}{\Delta V} = \dfrac{Q_t - Q_{t-1}}{V_t - V_{t-1}} \\ Q = \int I_c dt \end{cases} \tag{8}$$

Drawing the IC curves with the data of phase 1C in cell 7, which is presented in Fig. 4. However, Fig. 4 (a) covers that the original curves are not smooth, by using filtering technique to the original curves. The processed IC curve is shown in Fig. 4 (b).
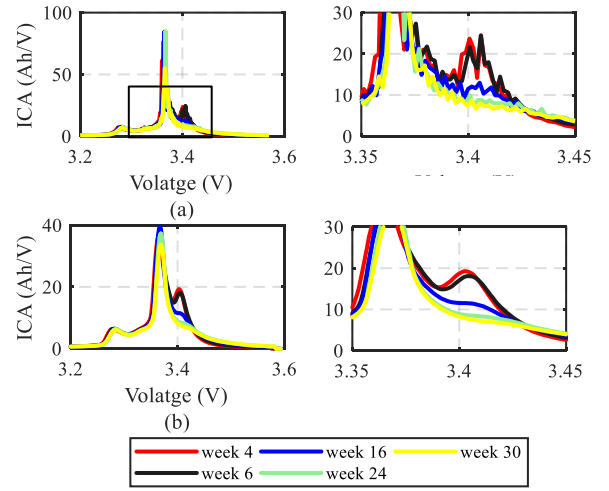

Fig. 4. The evolution of the IC curves for tested battery during the aging process,(a) original curves; (b) under filtering technique.

As shown in Fig. 5, the IC curve under the 1C phase and the 4C phase is plotted. In the 1C stage, it can be seen three obvious peaks, that are peak 1, peak 3 and peak 5, while also has two clear valleys, valley 2 and valley 4. In the 4C stage, it can only get one obvious peak, that is peak 6, and two inconspicuous valley, valley 7 and valley 8.

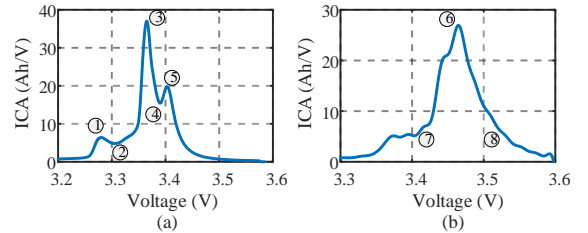Fig. 5. The characteristic of IC curves, (a) under 1C charging rate; (b) under 4C charging rate.

Fig. 6 shows the evolution of the IC curves for battery under different the operating condition. Based on Table III, we extracted the feature from the IC curves, such as the peak value, that is position of the peak and ICA value of the peak, the slope of the peak, the area of the peak and the same characteristic of valley.
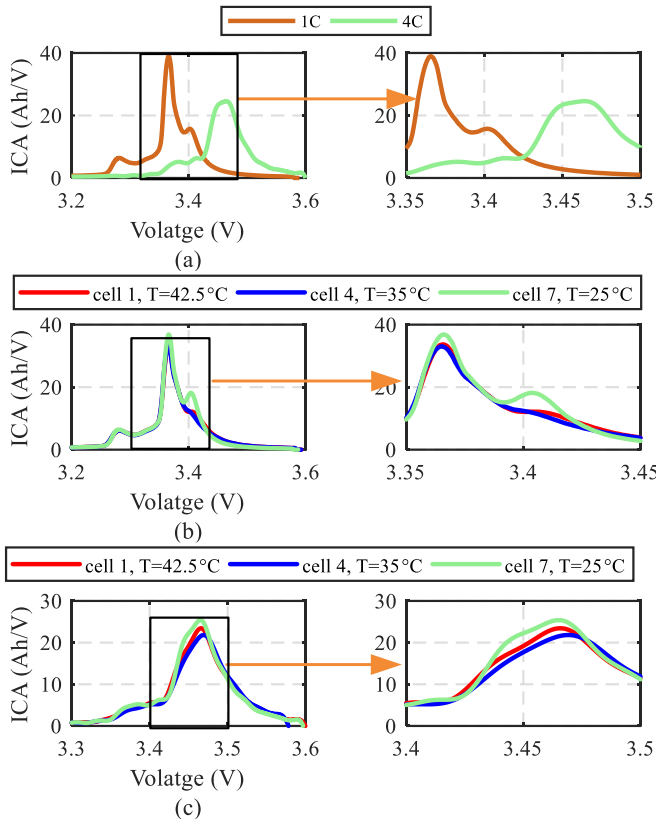
Fig. 6. The evolution of the IC curves for battery during the aging condition,(a) different charging rate; (b) different temperature with 1C; (c) different temperature with 4C.

TABLE III. THE DESCRIPTION FOR FEATURE PARAMETERS BASED ON IC CURVES

Feature
Peak
Right/Left Slope
Valley
Area

Based on above analysis, the focus is that to extract the health indicators from the IC curves. According to Fig. 5 and Fig. 6, the position of the peak/valley change significantly with changes in temperature, aging state, and charging rate. And the slope for the IC curves also has tilted change. Besides, the area near the peak/valley point is also considered. By using simple mathematical calculations, 31 features can be obtained, which shown in Fig. 7. Let the notation P to represent the peak, and V to represent the valley.
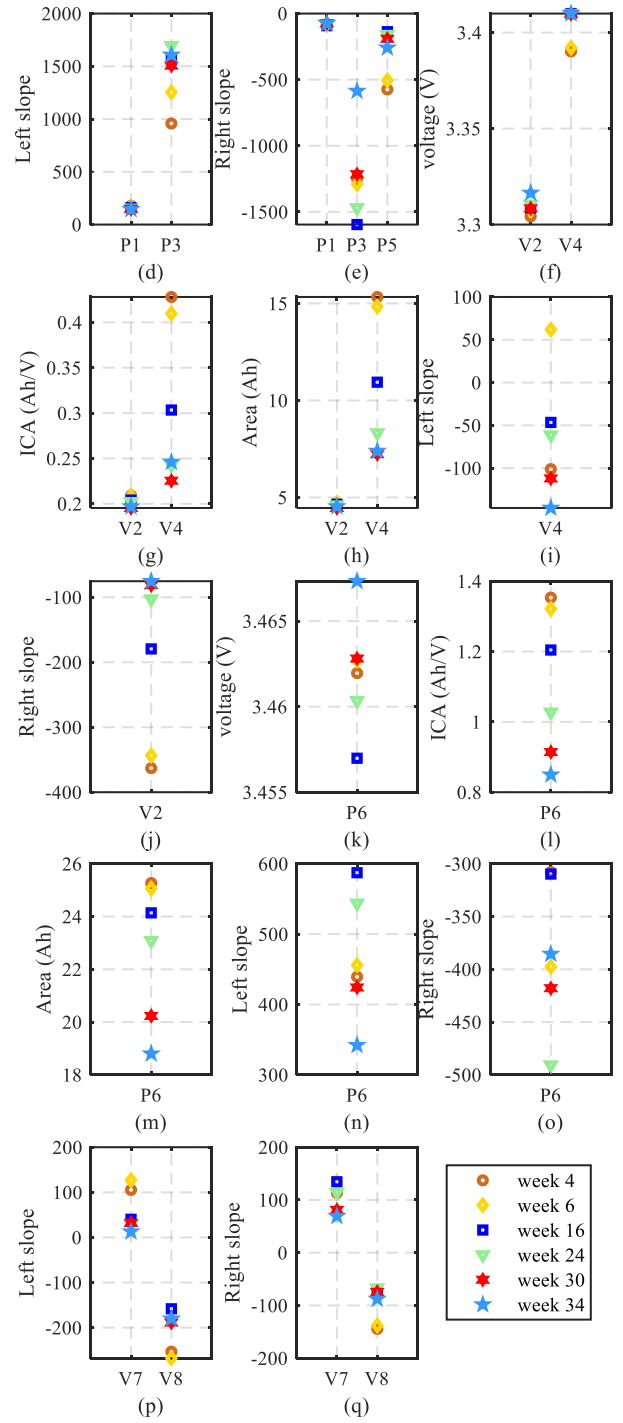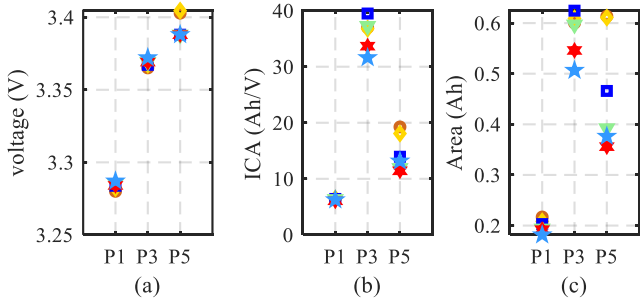

Fig. 7. The extracted features based on Cell 7.

V. ESTIMATED RESULTS AND DISCUSSION

In this part, we perform the effectiveness of the prognostic approach for estimating battery SOH. Eight cells are used for verification, while the verified methods under four types with different model input, type 1 is that using the health indicators with PCA processing; type 2 is that using all health indicators; type 3 is that using health indicators based on grey relation analysis, while the dimension equals to type 1; type 4 is that using health indicators based on randomly selected, while the dimension also equals to type 1. Besides, three types errors are chosen to measure the estimated accuracy, that are the mean absolute error, mean square error and root mean square error. Finally, by using the coefficient of determination R_2 to

analysis the efficiency of the estimate model, where the larger value of the R_2, the better the model fitting effect.

Based on Table II, it has three condition battery set, that is TC1, TC2 and TC3. Different from other verified method, which the training set and testing set are from the same cell. In this paper, we choose the cell A in mode I as the training set, while the cell B in mode I as the testing set. This verification method is more convincing, even in the absence of a battery data, it still has an accurate prediction effect on its own SOH.

Besides, in order to select the appropriate number of the health indicators, the accumulated variance contribution rate of components produced by PCA are presented in Table IV, which means that the accumulated variance contribution rate of component 1 to component 7 of eight cells have reached above 99%. Therefore, we choose these six components as the final model input.
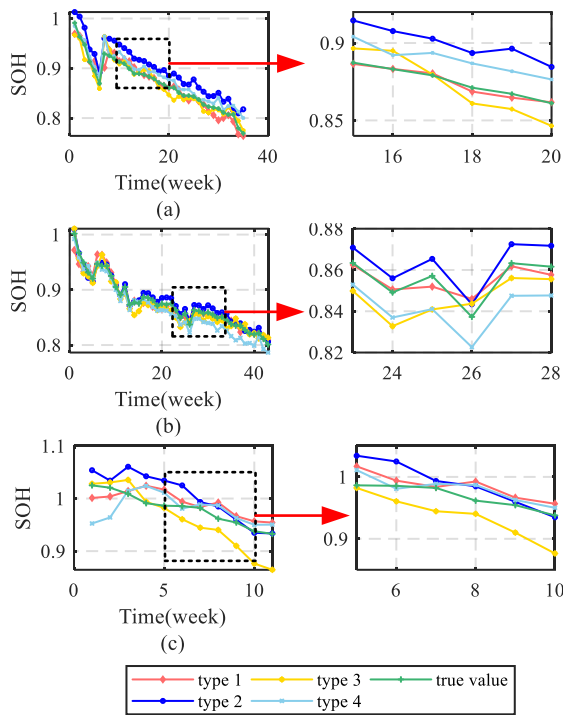

Fig. 8. The estimated result for (a) TC1, (b) TC2 and (c) TC3.

Fig. 8 shows the estimated result for TC1, TC2 and TC3 under four verified methods, which shows that the estimated result with type 1 can better track true value. Although, for mode 3, the estimated result with type 2 also can get accurate result, while the prediction accuracy based on type 1 is not much worse than it, this means can explain that even when all health indictors are used for training the model, the accuracy of the SOH estimation will not be improved.
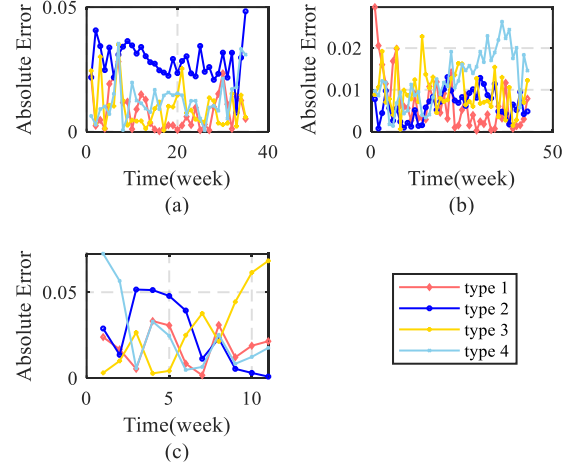
Fig. 9. The estimated absolute error for (a) TC1, (b) TC2 and (c) TC3.
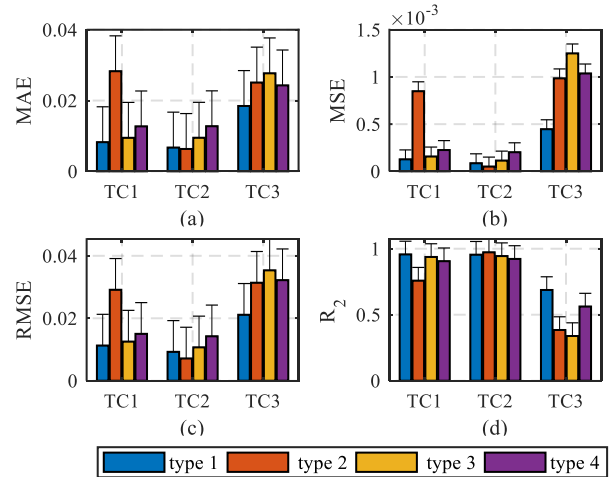
Fig. 10. Three errors and coefficient of determination R_2 for different type, (a) MAE, (b) MSE (c) RMSE (d) R_2.

TABLE IV. ACCUMULATED CONTRIBUTION RATE OF COMPONENTS.

		Cumulative contribution rate (100%)						
		component 1	component 2	component 3	component 4	component 5	component 6	component 7
TC1	cell 1	76.1%	21.37%	1.6%	0.4%	0.2%	8.27e-02%	2.61e-02%
	cell 2	79.4%	18.4%	1.1%	0.8%	0.2%	6.71e-02%	2.79e-02%
	cell 3	83.9%	13.9%	1.3%	0.5%	0.2%	9.93e-02%	5.38e-02%
TC2	cell 4	88.8%	8.9%	1.2%	0.5%	0.3%	0.2%	7.50e-02%
	cell 5	84.4%	12.4%	1.6%	1%	0.3%	0.2%	8.83e-02%
	cell 6	87.9%	9.4%	1.5%	0.7%	0.3%	0.1%	5.39e-02%
TC3	cell 7	50.7%	35.8%	7.8%	2.7%	1.6%	0.9%	0.3%
	cell 8	66.5%	25.5%	4.4%	2.3%	0.7%	0.2%	0.1%

Fig. 9 shows the estimated absolute error for TC1, TC2 and TC3 under four verified methods, which shows that the accuracy of SOH estimation under type 1 is higher than accuracy obtained for the other three types, and the estimated absolute error under type 1 is more stabler and smaller.

TABLE V. THE THREE ESTIMATED ERRORS FOR TC1 BASED ON DIFFERENT TYPE

Error	Type1	Type2	Type3	Type4
MAE	0.0082	0.0283	0.0095	0.0127
MSE	1.27e-04	8.49e-04	1.57-e04	2.25e-04
RMSE	0.0113	0.0291	0.0125	0.015

TABLE VI. THE THREE ESTIMATED ERRORS FOR TC2 BASED ON DIFFERENT TYPE

Error	Type1	Type2	Type3	Type4
MAE	0.0067	0.0063	0.0095	0.0127
MSE	8.65e-05	5.07e-05	1.14-e04	2.02e-04
RMSE	0.0093	0.0071	0.0107	0.0142

TABLE VII. THE THREE ESTIMATED ERRORS FOR TC3 BASED ON DIFFERENT TYPE

Error	Type1	Type2	Type3	Type4
MAE	0.0185	0.0251	0.0277	0.0243
MSE	4.46e-04	9.85e-04	0.0012	0.001
RMSE	0.0211	0.0314	0.0353	0.0322

Fig. 10 and Table V - Table VII show the three types errors and the coefficient of determination R_2 for TC1, TC2 and TC3 under four verified methods. For TC1, the MAE and RMSE using type1 SOH estimation is about 0.0082 and 0.0113, while the MSE is approximately 1.27-e04, which is slower than other three types. For TC2, the MAE and RMSE using type1 SOH estimation is about 0.0067 and 0.0093, while the MSE is approximately 8.56-e05. For TC3, the MAE and RMSE using type1 SOH estimation is about 0.0185 and 0.0211, while the MSE is approximately 4.46-e04. Through Fig. 10 (d), we find that the R_2 of the type 1 under three mode is higher than other three types, which means that based on type 1, it can get better battery SOH estimated results.

VI. CONCLUSION

This paper proposed the technique for estimating battery SOH based on ICA technique. By processing the battery raw data with ICA technique, 31 health indicators are extracted from IC curves. Besides, the PCA algorithm reduces the number of the model input so as to eliminating the negative effect of computational burden caused by multi-dimensional input; the results show that multi-dimensional input does not improve the prediction accuracy. In addition, the training and testing data set are from different data set, which is different from common verified method. Then, the mean square error

of the proposed method is less than 5.50e-04 for these three mode, which shows the proposed method has a high accuracy.

REFERENCES.

[1] X. Hu, L. Xu, X. Lin, M. Pecht, Battery lifetime prognostics, Joule 4 (2) (2020) 310 – 346.

[2] C. R. Lashway, O. A. Mohammed, Adaptive battery management and parameter estimation through physics-based modeling and experimental verification, IEEE Transactions on Transportation Electrification 2 (4) (2016) 454–464.

[3] T. Ashwin, Y. M. Chung, J. Wang, Capacity fade modelling of lithium-ion battery under cyclic loading conditions, Journal of Power Sources 328 (2016) 586 – 598.

[4] C. Liu, Y. Wang, Z. Chen, Degradation model and cycle life prediction for lithium-ion battery used in hybrid energy storage system, Energy 166 (2019) 796 – 806.

[5] M. Berecibar, F. Devriendt, M. Dubarry, I. Villarreal, N. Omar, W. Verbeke, J. V. Mierlo, Online state of health estimation on nmc cells based on predictive analytics, Journal of Power Sources 320 (2016) 239–250.

[6] B. Saha, K. Goebel, S. Poll, J. Christophersen, Prognostics methods for battery health monitoring using a bayesian framework, IEEE Transactions on Instrumentation and Measurement 58 (2) (2009) 291–296.

[7] X. Li, C. Yuan, Z. Wang, State of health estimation for li-ion battery via partial incremental capacity analysis based on support vector regression,Energy 203 (2020) 117852.

[8] D. Stroe, E. Schaltz, Lithium-ion battery state-of-health estimation using the incremental capacity analysis technique, IEEE Transactions on Industry Applications 56 (1) (2020) 678–685.

[9] M. Dubarry, V. Svoboda, R. Hwu, B. Y. Liaw, Incremental capacity analysis and close-to-equilibrium OCV measurements to quantify capacity fade in commercial rechargeable lithium batteries, Electrochemical and Solid-State Letters 9 (10) (2006) A454–A457.

[10] M. Berecibar, M. Dubarry, N. Omar, I. Villarreal, J. Van Mierlo, Degradation mechanism detection for nmc batteries based on incremental capacity curves, World Electric Vehicle Journal 8 (2) (2016) 350–361.

[11] A. Fly, R. Chen, Rate dependency of incremental capacity analysis (dq/dv) as a diagnostic tool for lithium-ion batteries, Journal of Energy Storage 29 (2020) 101329.

[12] X. Li, Z. Wang, L. Zhang, C. Zou, D. D. Dorrell, State-of-health estimation for li-ion batteries by combing the incremental capacity analysis method with grey relational analysis, Journal of Power Sources 410-411 (2019) 106–114.

[13] X. Li, Z. Wang, J. Yan, Prognostic health condition for lithium battery using the partial incremental capacity and gaussian process regression, Journal of Power Sources 421 (2019) 56 – 67.

[14] D. Yang, X. Zhang, R. Pan, Y. Wang, Z. Chen, A novel gaussia process regression model for state-of-health estimation of lithium-ion battery using charging curve, Journal of Power Sources 384 (2018) 387–395.

Convolutional LSTM Networks for Vibration-Based Defect Identification of the Composite Structure

Ruiheng Zhang
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
zrhtfsc2@hotmail.com

Libing Bai*
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
bailb991@163.com

Jie Zhang
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
zhj0619@uestc.edu.cn

Lulu Tian
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
tl.110119@163.com

Rui Li
Shanghai Space Propulsion
Technology Research Institute
Shanghai, China
22416812@qq.com

Zhen Liu
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
scdliu@uestc.edu.cn

Abstract—**The increasing need for reliable and sensitive detection methods have been brought about by the design of lighter, more flexible or complicated composite structures. In the field of vibration-based testing, many studies about the assessment of composite structures have been proposed. Artificial neural networks(ANNs) have offered a novel route to explain the relationship between vibration feature and defect status, and an increasing number of ANN frameworks have been designed for adhesive defect identification in the composite structure. This paper proposed a novel ANN framework based on Convolutional Long Short-Term Memory (ConvLSTM) in order to identify composite structure with adhesive defect. The proposed framework allowed pattern classification of the defect accordingly. To verify the feasibility of the proposed approach, two specimens with different levels of debonding were tested. The obtained results indicate that the classification accuracy strongly depended on the proper adaption of ConvLSTM**

Index Terms—*Artificial neural networks, adhesive defect identification, ConvLSTM.*

I. INTRODUCTION

Composite structures, such as steel-concrete structures, laminated plates, adhesive joints, etc., have been employed across a wide range of engineering fields, as they are often designed to satisfy specific requirements such as bearing capacity, durability, insulation, or corrosion resistance in these fields. In recent years, the increasing need for reliable and sensitive detection methods have been brought about by the design of lighter, more flexible or complicated composite structures, and many studies about the assessment of composite structures have been proposed in the field of vibration-based testing/measurement [1]. The dynamic properties of composite structures such as modulus of elasticity, natural frequency, stiffness, damping factor, etc. can be acquired from vibration responses due to the fact that they are, in general, susceptible to the action of dynamic force [2]. However, it has also been theoretically and practically proved that a local defect like delamination, and debonding in a reference structure can cause the reduction of local structural stiffness, changing the dynamic properties of the whole structure [3] and imposing a huge threat to the structural safety. Therefore, the detection for the composite structure defects is an emergent concern in applications and represents a more challenging task in consideration of the fact that the relationships between vibration signals and these defects are quite complicated or there is no known analytical solution at all [4].

Recently, artificial neural networks(ANNs) have presented a great advance in establishing the mapping relationship between vibration feature and defect status. An increasing number of ANN frameworks have been designed for adhesive defect identification in the composite structure. The mostly used ANN in adhesive defect identification was multi-layer backpropagation network (MLBPN) [5] [6]. Since the vibration signals measured in the time domain are composed of thousands of samples, the determination of suitable feature proxy as training data for ANN becomes a very important issue [7]. Although the aforementioned methods can identify adhesive defects in composite structures, most of them predominantly depend on the choice of feature extraction techniques, which usually require complex signal processing. Moreover, in ANN based method, a neural network was often used as the classifier for defect identification where its feature learning capability was constrained. There remains a need for a compact ANN framework for adhesive defect identification that includes both feature extraction and classification parts.

One may say that the convolutional neural network(CNN) is applicable for feature extraction because it can automatically extract high-level features from the raw dataset and preserve the spatial information [8]. However, in the field of vibration testing, signals collected from a sensor network, e.g. velocity, acceleration, and displacement are regarded as a spatiotemporal sequence, whereas CNN can not take the temporal correlations of the sequence into consideration. In

---

* Libing Bai is the corresponding author. (e-mail: bailb991@163.com).

contrast to CNN, Convolutional Long Short-Term Memory(ConvLSTM) is made up of convolution structures in both the input-tostate and state-to-state transitions [9], which makes it possible to capture spatiotemporal information from the input data, and produce more discriminative spatiotemporal feature representations. ConvLSTM has proven powerful for capturing spatiotemporal correlations from videos, weather maps, and traffic flows [10] [11] [12], but few studies have focused on vibration signals measured from a sensor network. Inspired by the above work, this paper proposed a novel ANN framework based on ConvLSTM for adhesive defect identification.

The main objective is to identify composite structure with adhesive defect by processing the vibration response acquired from a network of transceivers deployed on the structure. The proposed framework allowed pattern classification of the defect. To verify the feasibility of the proposed approach, two specimens with and without debonding defect were tested. The obtained results indicate that the classification accuracy strongly depended on the proper adaption of ConvLSTM. The main contributions of this paper are listed as follows:

(1) Existing methods mainly use ANN as a classifier in vibration testing. In many cases, the raw vibration signals are preprocessed by advanced signal processing techniques to extract distinguishable features. In this paper, in order to better utilize the feature extraction capability of ANN, ConvLSTM, which can retain the spatiotemporal information of vibration signals, is adopted to set up a compact network for adhesive debonding identification. Preprocessing the input data is simplified since the feature extraction task is implemented in the proposed network.

(2) Extensive experiments are carried out to analyze and verify the proposed model. In this paper, to analyze the feature extracting performance of the proposed network, we collected intermediate outputs of the proposed network under different configurations. The ConvLSTM layers' output difference is observed between debonding feature map and intact feature map, which indicate that ConvLSTM layers in the proposed network have managed to extract features from vibration signals. Furthermore, we compare the performance of the proposed network with traditional LSTM network. Without spatial information, the prediction accuracy of LSTM network is noticeably lower than that of the proposed network.

II. METHODOLOGY

Assume the vibration response of a reference structure in the time domain is collected from a sensor network mounted on the structure surface. It can be denoted as $X = \{X_{ij}(t)\}$, where $X_{ij}(t)$ is the vibration response collected from the sensor located in (i,j) at time t. As shown in [13], the mapping relationship between the excitation load and the response can be denoted as:

$$X = M(f(t)) \tag{1}$$

Where M is the mapping function between the excitation load f and the vibration response. It is worth noting that, in general, the relationship between the input and output of either linear or nonlinear structures can be described by a mapping function.
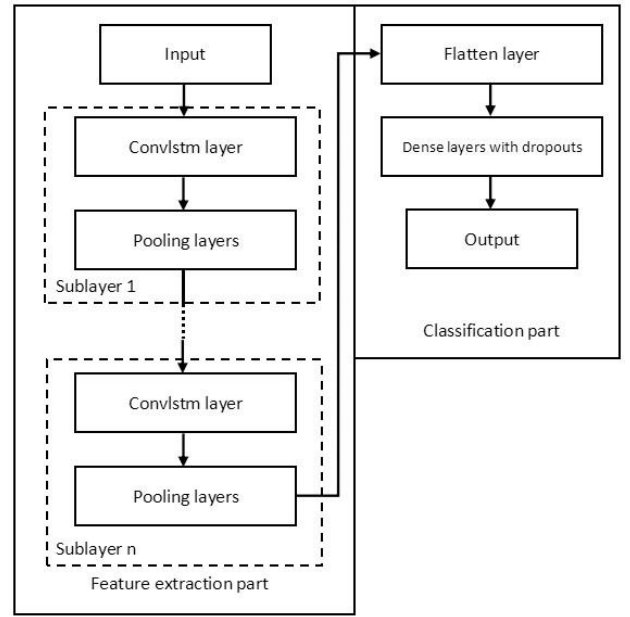

Fig. 1. The Debonding Identification Neural Network

Moreover, the mapping function is solely determined by the dynamic properties of the structure and when the load is invariant, the vibration response varies over the measured structures. By labeling vibration responses from different specimens as $Y = \{k\}$, (1) can be rewritten as:

$$X_k = M_k(f(t)) \tag{2}$$

As illustrated in fig. 1, the proposed neural network is mainly composed of two parts: feature learning and classification. The core structure of feature learning part consists of ConvLSTM layers and pooling layers and learns spatiotemporal features from the vibration response. The classification part uses dense layers as the classifier for the final debonding recognition. The key equations of the ConvLSTM cell are formulated as:

$$f_t = \sigma(w_f * [h_{t-1}, x_t] + b_f) \tag{3}$$

$$I_t = \sigma(w_i * [h_{t-1}, x_t] + b_i) \tag{4}$$

$$C_t \sim = tanh(w_c * [h_{t-1}, x_t] + b_c) \tag{5}$$

$$C_t = f_t * C_{t-1} + i_t * C_t \sim \tag{6}$$

$$o_t = \sigma(w_o * [h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = o_t * tanh(C_t) \tag{8}$$

Where sigma is the logistic sigmoid function, f_t, i_t, $C_t \sim$, C_t, o_t and h_t are respectively the forget gate, input gate, cell state candidate, cell state, output gate and hidden state vectors of the same size, w_f, w_i, w_c and w_o are weights of convolution windows, and b_f, b_i, b_c and b_o are biases. The pointwise product of f_t and C_{t-1} in (6) means that some information from the previous cell state(C_{t-1}) is removed by f_t. The elements in f_t are numbers between 0 and 1, where a 1 represents keeping all previous information while a 0 represents forgetting all previous information. Similarly, the pointwise product of i_t and C_t in (6) indicate that some information from candidate cell state($C_t \sim$) is selected by i_t. The size of vibration response

or the input of ConvLSTM layer X can be denoted as $M{\times}N$, and hidden state or the output of ConvLSTM layer H can be seen as the hidden representations of vibration response X with size $M{\times}N{\times}F$(padding=same), where F is the number of convolution windows. From (3)-(8), it is known that spatial features of X are extracted by these windows and temporal features are stored in cell state C. H, C separated into F components can be represented as:

$$H = c_h(H_1,...,H_i,...,H_F) \qquad (9)$$

$$C = c_h(C_1,....,C_i,....,C_F) \qquad (10)$$

Where c_h is horizontal concatenate function, C_i denotes the temporal feature of X filtered by the *ith* window, and H_i denotes the spatio-temporal feature of X filtered by the *ith* window. To reduce computational complexity, two types of pooling layers are used to shrink the size of H. An average pooling layer is used to reduce the temporal size of hidden states, and a max pooling layer is used to reduce spatial size. Depending on the complexity of the input data, multiple layers that comprise a ConvLSTM sublayer(a constituent ConvLSTM layer), an average pooling sublayer(a constituent average pooling layer), and a max pooling sublayer(a constituent max pooling layer), can be stacked to achieve the desired feature representation for the input data. Therefore, the output of the feature learning part stacking N layers can be represented as:

$$O = T_n(X) \qquad (11)$$

Where T_n is the mapping function between X and the *nth* feature extraction layer's output. The conditional probability distribution of each composite structure sample predicted by classification part is represented as:

$$Q_i(y = i|fl(O)) \qquad (12)$$

Where fl is the flatten function, Q_i is the probability value of the *ith* vibration response x_i can be obtained from the
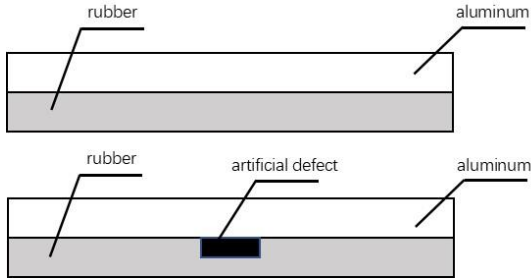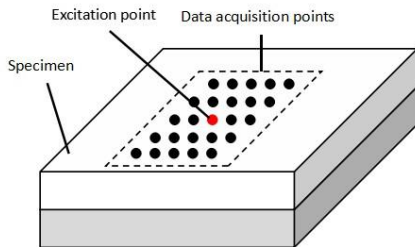

Fig. 2. Specimens for Testing

Fig. 3. Experimental Schematic

neural network, and i indicates the labels of the composite structure samples. The loss function of the whole network can be described as:

$$Loss = \sum_{i=1}^{N} P_i * log(Q_i) \qquad (13)$$

where P_i is the true probability distribution of labels and (13) is used to learn the network weights and biases, and is optimized by the adaptive moment estimation algorithm(Adam) in this work.

III. EXPERIMENT SETUO AND SAMPLES

The specimens for testing are shown in Fig. 2, and the dimension of these two plates is $30{\times}30{\times}3mm$. The composite plates were composed of two parts, an aluminium and a rubber, which was bonded by epoxy resin.

The first specimen was an intact specimen with a full connection between both parts. The second specimen had an artificial debonding of size $2{\times}2mm$. Fig. 3 is the experimental schematic diagram of the test setup. For each measurement, two ends of the specimen are fixed by metal strips. The vibrator applied a continuous sinusoidal load with a frequency of 5kHz on the surface of the specimen at excitation points. The accelerometer network of 8×8 was mounted on the top of the specimen to collect the vibration response. In total, two groups of vibration responses of size $100000 \times 8 \times 8$ from excitation point 1 have been used to testify the proposed identification network.

Three steps were conducted before training the proposed identification network:

(1) The raw data of vibration responses were denoised by the bandpass filter.

(2) The denoised data were symmetrically dilated in spatial dimensions since the spatial size of the data was small. It is worth noting that when the spatial size of the data is large
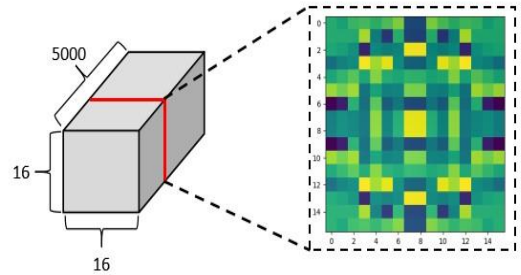

Fig. 4. The Training Sample Data Block

(e.g. large scale data from a full-field scanning vibrometer), there is no need for spatial dilation;

(3) The training segments of size $5000 \times 16 \times 16$ were randomly picked from the dilated data, which means each segment has different starting times. This is mainly for testing the temporal feature learning capability of our network. For each group of vibration response, we picked 200 segments. Therefore, the final training dataset for each group is a $200 \times 5000 \times 16 \times 16$ tensor. Suppose the batch size is 1, then one training sample is a $5000{\times}16{\times}16$ data block as shown in Fig.4.

IV. RESULTS AND DISCUSSION

We extensively evaluate the proposed network and the learnt spatiotemporal features under various settings for the task of debonding identification.

To quantitatively and qualitatively analyze the feature extraction capability of the proposed model, several ConvLSTM based models with different structures and

configurations were analyzed. Furthermore, we compare the proposed network with the LSTM network using the same dataset. It is worth noting that the training sample of the LSTM network does not need dilation and is dimensionally reduced to 2D of size 5000 × 64(5000 timesteps and 64features for each step).

TABLE I. GODEL ACCURACY AND CONFIGURATION

Models	Accuracy	Configuration
1	0.5	ConvLSTM 1(filters=40 kernelsize=(6,6))
		ConvLSTM 2(filters=20 kernelsize=(3,3))
2	0.64	ConvLSTM 1(filters=40 kernelsize=(6,6))
		MaxPooling 1(2,2,2)
		ConvLSTM 2(filters=20 kernelsize=(3,3))
		MaxPooling 2(2,2,2)
3	0.99	ConvLSTM 1(filters=40 kernelsize=(6,6))
		MaxPooling 1(5,3,3)
		ConvLSTM 2(filters=20 kernelsize=(3,3))
		MaxPooling 2(3,2,2)

As illustrated in table 1, the main difference between these three models is their different configurations of ConvLSTM layers and pooling layers and the prediction accuracy of each model varies under different configurations.

To further investigate the behavior of these models, we analyzed two intermediate outputs for each model. Intermediate output#1 is the output of feature extraction part of the model and intermediate output#2 is the output of the classification part.
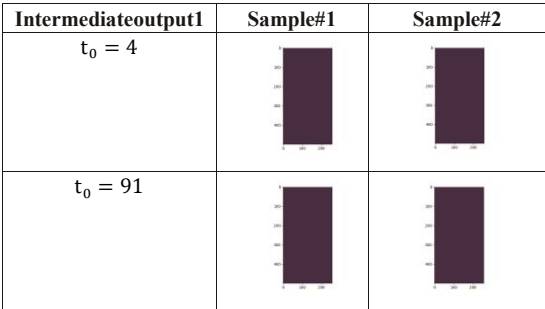
Intermediateoutput1	Sample#1	Sample#2
$t_0 = 4$		
$t_0 = 91$		

Fig. 5. Intermediate Output 1 of Model 1

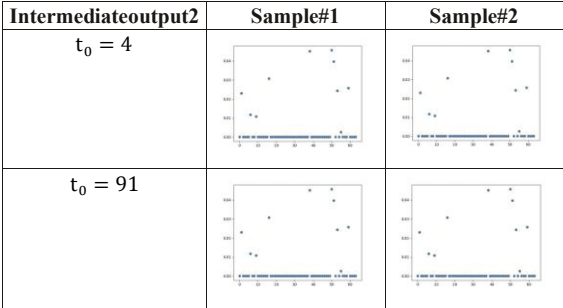| Intermediateoutput2	Sample#1	Sample#2
$t_0 = 4$		
$t_0 = 91$		

Fig. 6. Intermediate Output 2 of Model 1

As shown in Fig.5 and Fig. 7, where the vertical category includes feature maps of one sample with different starting times and the horizontal category includes feature maps of different samples' training data, the intermediate output#1 feature maps of low accuracy models(model1,model2) are almost indistinguishable in both horizontal category and vertical category.

Similarly, As shown in Fig.6 and Fig.8, the intermediate output#2 feature maps of low accuracy model are indistinguishable in both horizontal and vertical categories.

By systematically adjusting the configuration of the Convlstm model, we obtained a high accuracy model(model3). However, Neural network configuring is beyond the scope of this paper and a further discussion is unnecessary.
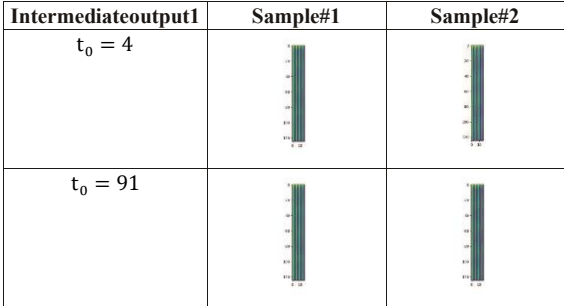
Intermediateoutput1	Sample#1	Sample#2
$t_0 = 4$		
$t_0 = 91$		

Fig. 7. Intermediate Output 1 of Model 2

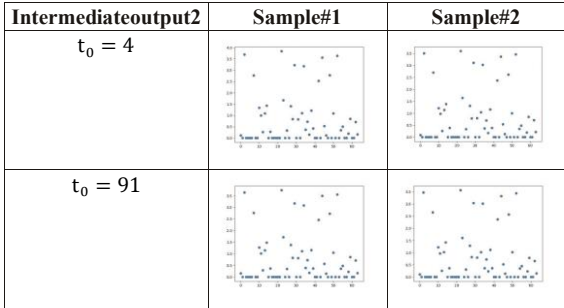| Intermediateoutput2	Sample#1	Sample#2
$t_0 = 4$		
$t_0 = 91$		

Fig. 8. Intermediate Output 2 of Model 2

As shown in Fig.9 and Fig.10, both intermediate output#1 and intermediate output#2 show some sort of differences in

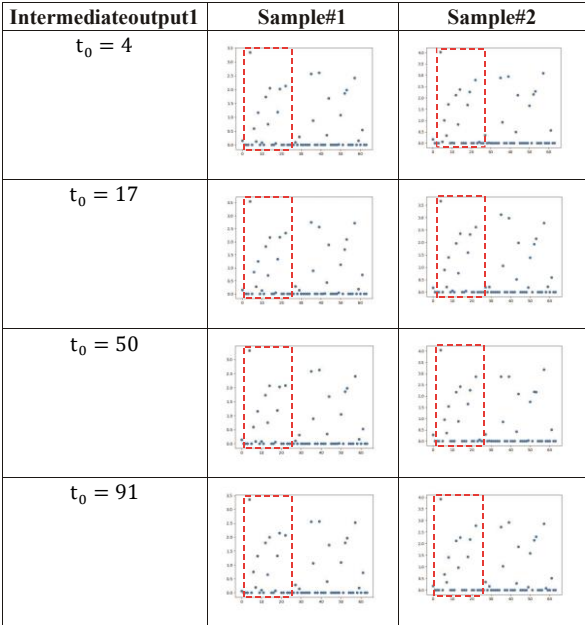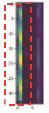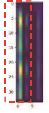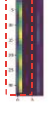| Intermediateoutput1	Sample#1	Sample#2
$t_0 = 4$		
$t_0 = 17$		
$t_0 = 50$		
$t_0 = 91$		

Fig. 9. Intermediate Output 1 of Model 3
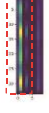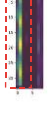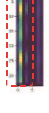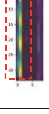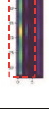
Intermediateoutput2	Sample#1	Sample#2
$t_0 = 4$		
$t_0 = 17$		
$t_0 = 50$		
$t_0 = 91$		

Fig. 10. Intermediate Output 2 of Model 3

rows and columns, but the more obvious differences occur horizontally. Compared to Sample#2's feature maps of intermediate output#1, Sample#1's feature maps have more compactly distributed yellow areas. Similarly, the feature maps of Sample#1 and Sample#2 have different distributions, especially inside the red dotted box.

As illustrated in table 2, the ConvLSTM model outperforms the LSTM model in terms of accuracy. In addition, the trainable parameters needed for the ConvLSTM model(544395) is much lower than that for the LSTM model(12872137), which suggest that deep feature extraction alone is not competent to the task of spatiotemporal dataset identification.

TABLE II.   MODEL ACCURACY AND CONFIGURATION

Models	Accuracy	Configuration
LSTM	0.6	LSTM 1(unit=80)
		LSTM 2(unit=40)
		LSTM 3(unit=20)
		LSTM 4(unit=10)
ConvLSTM	0.99	ConvLSTM 1(filters=40kernelsize=(6,6))
		MaxPooling 1(5,3,3)
		ConvLSTM 2(filters=20 kernelsize=(3,3))
		MaxPooling 2(3,2,2)

As shown in fig.11 and fig.12, there is a distinct difference between the ConvLSTM intermediate outputs and the LSTM intermediate outputs. The original intermediate feature maps of the LSTM model are 10000×10. For better illustration, we change these maps' size as 1000×100. For better illustration, we change these maps' size as 1000×100. The ConvLSTM intermediate outputs of sample#1 and sample#2 while the intermediate outputs of the LSTM model between sample#1 and sample#2 are almost unable to discern.

What is more, the size of the ConvLSTM is much smaller than that of the LSTM model, which is mainly on account of the pooling operation.

Through the above investigation, we can find that the ConvLSTM model can extract distinguishable spatiotemporal features after proper configuration whereas the LSTM model given greater computation power(more trainable parameters) is not applicable to identify spatiotemporal dataset in this work.

t_0=4	Sample#1	Sample#2
LSTM		
ConvLSTM		

Fig. 11. Intermediate Outputs of LSTM and ConvLSTM Model at t_0=4

t_0=91	Sample#1	Sample#2
LSTM		
ConvLSTM		

Fig. 12. Intermediate Outputs of LSTM and ConvLSTM Model at t_0=91

V. CONCLUSION

In this paper, a debonding identification neural network based on ConvLSTM has been proposed to identify composite structure with debonding defect. This proposed method discusses the advantages of ConvLSTM in vibration-based testing. It is helpful for the future work to use the ConvLSTM for feature extraction and classification. The proposed network can reserve spatial and temporal information of vibration responses and result in better classification performance. This paper only presents the identification network for periodic vibration response. The state-of-the-art advances of convolution and recurrent networks can be further utilized in a compact form to identify more complex vibration responses(e.g. impulse and multi-frequency responses) from the damaged composite structure.

REFERENCES

[1]   R. Medeiros, G.S.C. Souza, D.E.T. Marques, F.R. Flor and V. Tita, Vibration-based structural monitoring of bi-clamped metal-composite bonded joint: Experimental and numerical analyses, *Journal of Adhesion*, vol. 00, pp. 1-27, 2020.

[2]   G. Paolo L and F. Vittorio, Applied Structural and Mechanical Vibrations. *Crc Press*, vol. 23(1), pp. 146-146, 2013.

[3]   Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj and D.J. Inman, Real-time vibration-based structural damage detection using onedimensional convolutional neural networks, *Journal of Sound and Vibration*, vol. 388, pp. 154-170, 2017

[4] M.S. Hossain, O.Z. Chao, Z. Ismail, S. Noroozi and S.Y. Khoo, Artificial neural networks for vibration based inverse parametric identifications: A review. *Applied Soft Computing Journal*, vol. 52, pp. 203219, 2017.

[5] S. Farhangdoust, S. Tashakori, A. Baghalian, A. Mehrabi and I.N. Tansel, Prediction of damage location in composite plates using artificial neural network modeling, *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems* 2019, pp. 20, 2019.

[6] R. Zhou and N. Li, Identification of adhesive quality of multi-interface adhesive members, *Proceedings - 2017 10th International Symposium on Computational Intelligence and Design, ISCID 2017*, vol. 1, pp. 228-231, 2017.

[7] Y.Y. Liu, Y.F. Ju, C.D. Duan and X.F. Zhao, Structure damage diagnosis using neural network and feature fusion, *Engineering Applications of Artificial Intelligence*, vol. 24(1), pp. 87-92, 2011.

[8] S.H. Feng, J.Y. Xu and H.B. Shen, Chapter Seven - Artificial intelligence in bioinformatics: Automated methodology development for protein residue contact map prediction In Biomedical Engineering, Biomedical Information Technology (Second Edition), *Academic Press*, pp. 217-237, 2020.

[9] X. Shi, Z. Chen, H. Wang, D.Y. Yeung, W.K. Wong and W.C. Woo, Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems*, pp. 802-810, 2015.

[10] Y. Li, H. Xu, M. Bian and J. Xiao, Attention based CNN-convlstm for pedestrian attribute recognition. *Sensors (Switzerland)*, vol. 20(3), pp. 1-14, 2020.

[11] Z. Sun and M. Zhao. Short-Term Wind Power Forecasting Based on VMD Decomposition, ConvLSTM Networks and Error Analysis, *IEEE Access*, vol. 8, pp. 134422-134434, 2020.

[12] Z.N. Yuan, X. Zhou and T.B. Yang, Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data. *In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 18)*, pp. 984-992, 2018.

[13] J.M. Zhou, L. Dong, W. Guan and J. Yan, Impact load identification of nonlinear structures using deep Recurrent Neural Network. *Mechanical Systems and Signal Processing*, vol. 133(28), pp. 106292-106292, 2019.

Study on the Lowest Spatial Resolution of Magnetic Flux Leakage Testing for Weld Cracks

Chunrui Feng
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
rui2157@outlook.com

Zhen Zhang
Shanghai Space Propulsion Technology
Research Institute
Shanghai, China
morrainny@163.com

Libing Bai*
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
bailb991@163.com

Lulu Tian
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
tl.110119@163.com

Jie Zhang
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
zhj06_19@uestc.edu.cn

Yuhua Cheng
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
yhcheng@uestc.edu.cn

Abstract—In this study, the analysis of the lowest spatial resolution for measuring weld cracks based on the magnetic flux leakage (MFL) test method was developed. MFL detection is a widely used method of pipeline nondestructive testing. The spatial resolution of the MFL sensors influences the data amount and detection accuracy. The magneto-optical detection technique can be used to solve the problem of insufficient spatial resolution, but the amount of data is very large and redundant. Using the lowest spatial resolution with ensuring accuracy can reduce the amount of data and decrease the time cost of analyzing data. A girth weld crack is a common and complex defect that occurs in pipelines, and high spatial resolution is required to ensure its detection accuracy. In this study, the MFL of weld cracks and its influencing factors were investigated. Based on the frequency domain analysis, the lowest spatial resolution was obtained, which was verified by experiments.

Keywords—magnetic flux leakage, weld cracks, spatial resolution

I. INTRODUCTION

Oil and gas pipelines are among the most crucial energy-transportation methods worldwide. Oil pipelines need to be inspected and maintained regularly, which involves a large market [1,2]. The magnetic flux leakage (MFL) method is a widely used nondestructive test technique, and it has the advantages of no coupling, low environmental requirements, and low costs [3]. In pipeline nondestructive testing, the MFL test method is commonly used [4]. The detector called pipeline pig is located inside the pipeline and propelled by the pressure of the fluid, as shown in Fig.1.
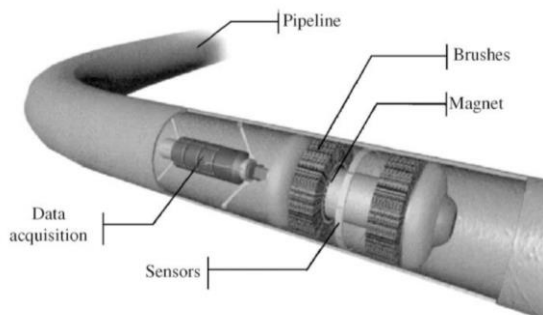
Fig.1. Pipeline pig [5]

The MFL test method is based on the measurement of the magnetic field near the surface of a specimen. The magnetic resistance of a defect is weaker than that of a normal location, which causes the magnetic flux in the specimen to leak into the air near the surface. The location of the defect can be determined by detecting the magnetic leakage field using magnetic sensors. The traditional pipeline detection method allows the pipeline pig to move along the pipeline with Hall sensors to measure the magnetic field near the inner pipe surface at short distances. For a single Hall sensor, the inspection result is a set of discrete signals along the axial direction of the pipeline. The number of checkpoints per unit distance represents the spatial resolution of the sensor.
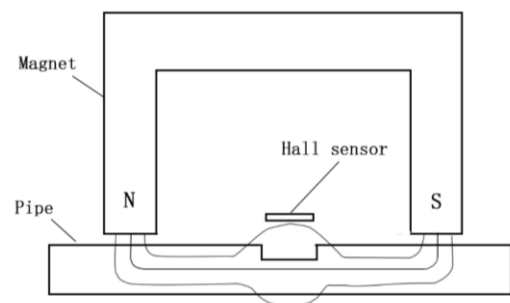
Fig.2. Schematic of MFL test instrument

One of the problems of the traditional MFL test technique is low detection accuracy [6]. The spatial resolution is restricted by the Hall sensor volume and pipeline pig speed. For some complex defect signals, vital information may be missed owing to insufficient spatial resolution. This issue can be solved by using the sensitive magneto-optical imaging (MOI) technique [7-10]. The MOI method is a type of detection technology based on the Faraday optical rotation effect shown in Fig.3, which can transform the signal of the magnetic field intensity into an optical intensity signal using polarizers. As shown in Fig.4, the MOI method can be applied to detect the magnetic field component perpendicular to the sample surface.

As a visual nondestructive test technique, the spatial resolution of the MOI can, theoretically, be equal to that of camera pixels. However, the MOI method has the challenge of receiving numerous information. The length of an oil and gas pipeline can extend up to several thousand kilometers.

Because of the high-precision camera, the amount of data is very large. Therefore, it is necessary to investigate how to eliminate redundant data and retain essential information.
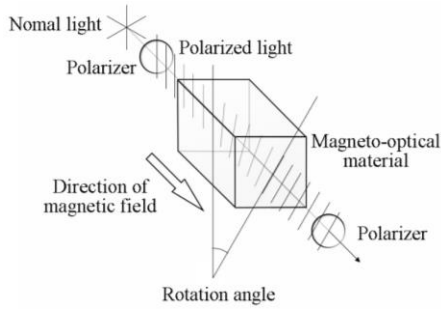
Fig. 3. Schematic of Faraday optical rotation effect
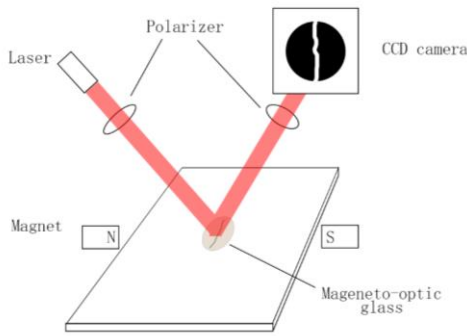
Fig. 4. Schematic of MOI method

The key to solving the above problem is adopting the lowest spatial resolution to ensure accuracy. Taking a single Hall sensor on pipeline pigs as an example, the sensor measures the value of the magnetic field at intervals along the axial direction of the pipeline. The detection result has a mapping relationship with the axial displacement. For the magneto-optic system, the pixels on the same line along the pipeline axis are equivalent to the detection signal of a single Hall sensor with a significantly higher detection density. The format of the detection results is the same as that of the time-domain signal. Therefore, the detection result can be analyzed in the frequency domain, which is similar to a time-domain signal.

The pipeline girth weld is highly prone to defects in pipelines, and it is a common type of complex signal interference in pipeline detection. It requires higher detection accuracy and magnetic spatial resolution for distinguishing weld defects. The most common defects can be detected if weld defects can be detected. In this study, the influence factors of the weld crack leakage magnetic field were analyzed. The lowest spatial resolution for detecting weld defects was determined through finite element simulation and frequency-domain analysis.

II. ANALYSIS OF MAGNETIC FLUX LEAKAGE OF WELD CRACKS

We aim to determine the lowest spatial resolution that ensures accuracy detection. Therefore, it is necessary to analyze the main factors that influence the magnetic field leakage of weld cracks. The variation of the signal waveform was determined to ensure that the vital data of the magnetic field leakage would not lost in any case.

The girth weld is the joint at two sections of the pipeline. At high temperatures, the edges of the pipe and welding rod are melted together to form a single unit, and the weld is formed after cooling. Compared to the normal position, the electromagnetic characteristics of the weld are different. The geometric shape of the weld is different from that of the normal part, and convex reinforcement is formed on the surface. In addition, owing to the process of solder filling and heat treatment, the permeability distribution of the weld is different from that of the normal position. Reinforcement and permeability are the two main factors that influenced the magnetic field leakage of the weld.

The magnetic field component perpendicular to the pipe surface is defined as the radial magnetic field. When cracks appear, the crack depth is the main factor that influence the radial magnetic field, while the crack width hardly influence the radial magnetic field. The detection-liftoff value is the most significant factor that influence the test results in all cases.

A. Effects of Crack Depth and Crack Width

When the liftoff value is 3 mm, the effects of crack depth and crack width on the plane without weld are shown in Fig.5 and Fig.6. The positive direction of the magnetic field is vertically upward. It can be seen that the deeper the depth, the stronger the radial magnetic field. Moreover, the influence of crack width can be ignored.
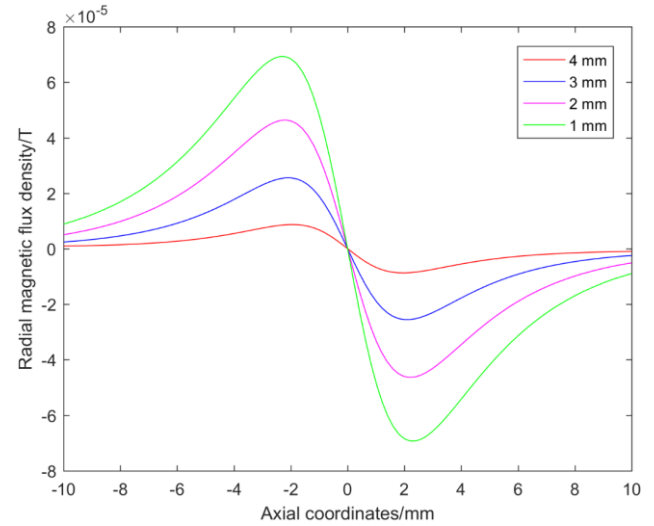

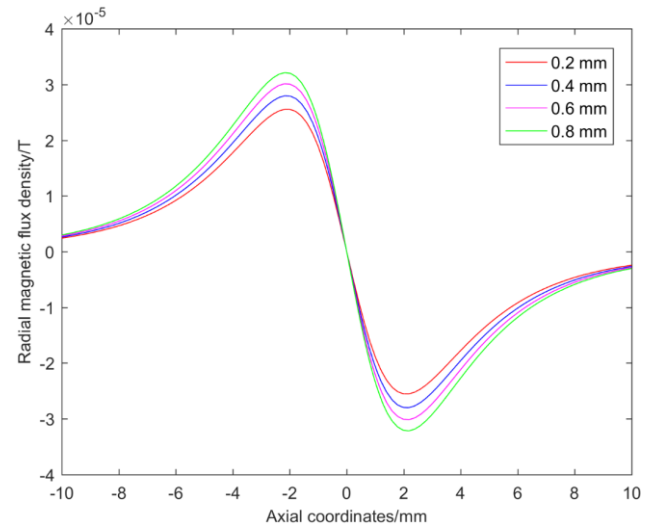Fig.5. Effect of depth on radial magnetic flux density

Fig.6. Effect of width on radial magnetic flux density

B. Effect of Reinforcement Height

The reinforcement is a convex surface, which increases the cross-sectional area and decreases the magnetic resistance. The reinforcement causes the magnetic flux in the air to enter inside the metal, and this creates a radial leakage magnetic field opposite the crack direction.
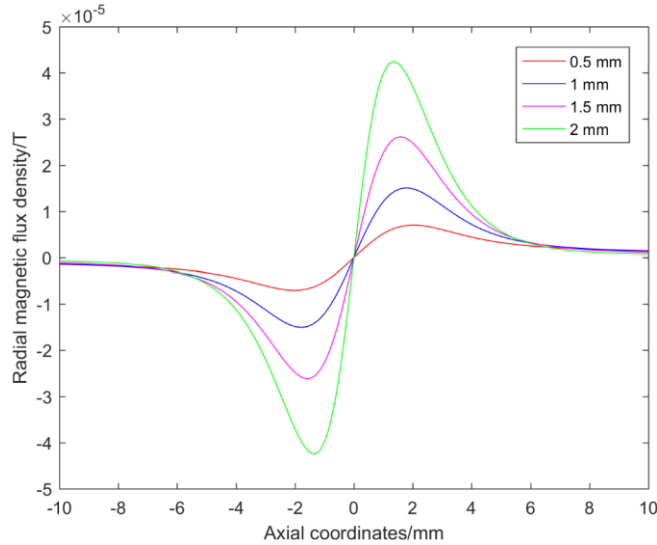

Fig.7. Effect of reinforcement height on radial magnetic flux density

The influence of different reinforcement heights on the radial magnetic field is obtained through finite element simulation. Fig.7 shows the radial magnetic field when the liftoff value is 3 mm; the abscissa is the axial displacement, and the positive direction of the magnetic field is vertically upward. It can be seen from the results that the higher the reinforcement height, the stronger the radial magnetic field. However, the radial magnetic flux leakage generated by the reinforcement is opposite to that of the crack. The higher the reinforcement height, the more disturbed the leakage magnetic field of the crack.

C. Effect of Permeability Distribution

The welding rod and inclusions modified the chemical composition of the weld in welding process. And the regional segregation effect results in the concentration of impurities to the weld center. The closer to the weld center, the more severe the lattice distortion. The permeability distribution of a weld could be simplified into three parts: base metal, transition zone, and central zone. The effect of the permeability distribution on the leakage magnetic field could be simulated by setting a decreasing permeability gradient.

The influence of the permeability distribution was determined via finite element simulation. Fig.8 shows the radial magnetic field with a liftoff value of 3 mm. The horizontal axis indicates the axial displacement, and the vertical direction shows the positive direction of the radial magnetic field. The three numbers represented the permeability, transition zone, and central zone of the base metal, and the results shows that the higher the permeability gradient, the more significant the magnetic leakage. In addition, it enhance the edge of the crack leakage magnetic field.
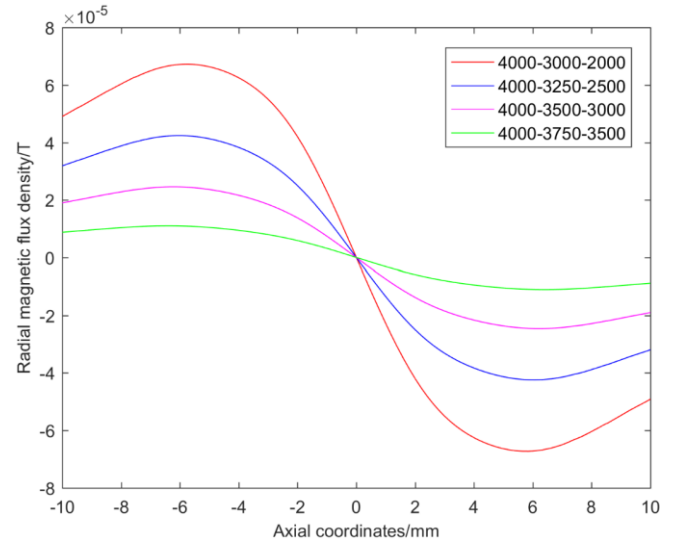
Fig.8. Effect of permeability distribution on radial magnetic flux density

D. Effect of Liftoff

The liftoff value directly influenced the test results, but it impossible to adhere to the specimen entirely during the actual test. Fig.9 shows that when the crack depth is 2 mm, the width, reinforcement height, and permeability distribution are 0.2 mm, 1 mm, and 4000-3500-3000, respectively. The liftoff value significantly influenced the radial magnetic field detection value. It can be seen that the liftoff value has a great influence on the detection results. The higher the liftoff value, the smoother the signal waveform.
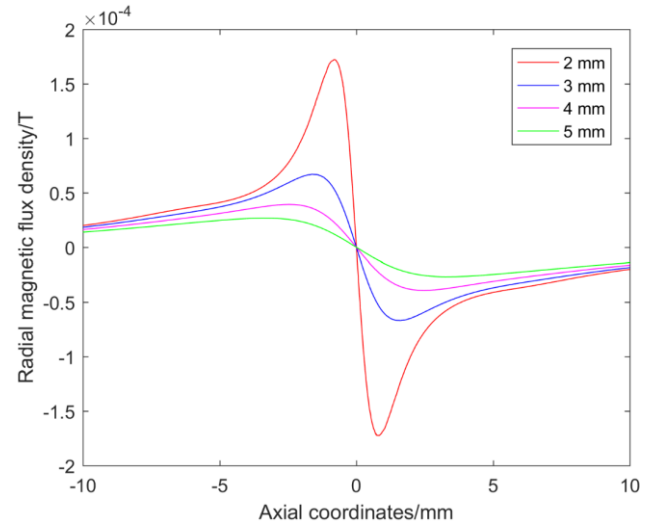
Fig.9. Effect of liftoff on radial magnetic flux density

III. FREQUENCY DOMAIN ANALYSIS

A mapping relationship between the axial displacement and radial magnetic field exists. The waveform of the radial magnetic field intensity can be plotted by taking the axial displacement as the abscissa. Therefore, it can be analogized with the time-domain signal and processed using the frequency-domain analysis method. The challenge of determining the lowest spatial resolution is changed to finding the lowest sampling rate of the signal waveform, which could be easily solved using the sampling theorem. The signal waveform of weld crack has the characteristics of smooth edge part and steep middle part, which could be

predicted as a low-pass signal after frequency domain transformation.

The signal waveform of a typical weld crack was transformed via fast Fourier transform. The result is shown in the Fig.10. The frequency-domain signal was concentrated at the low frequency. If the low-pass cutoff frequency is known, the lowest sampling rate could be obtained using the Nyquist sampling theorem.

(a)   (b)
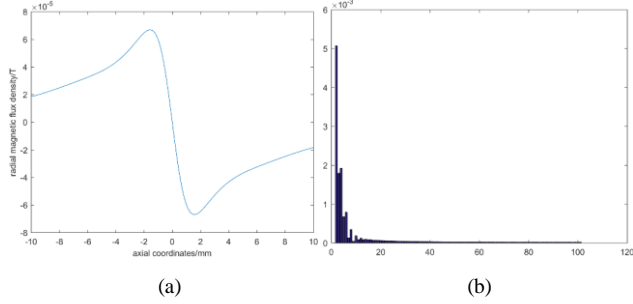
Fig.10. (a) Signal of typical weld crack (b) Frequency domain

The low-pass cutoff frequency was determined using the Pearson correlation coefficient before and after low-pass filtering. The Pearson correlation coefficient is an index used for evaluating the similarity of two signal waveforms. The closer the value is to 1, the more familiar the two waveforms are. The threshold of the Pearson correlation coefficient was set to 0.99, and the low-pass cutoff frequency was the lowest frequency that caused the correlation coefficient of the waveforms before and after filtering to be higher than 0.99. The number of checkpoints within 1 mm displacement was defined as the sampling rate. For these parameters, the low-pass cutoff frequency of the above signal is 0.3 Hz. According to the Nyquist sampling theorem, the lowest sampling frequency is 0.6 Hz. The lowest spatial resolution is 0.6 times per mm, and the maximum distance between two detection points is 1.6 mm. The reconstruction signal based on the 0.6 Hz sampling rate is shown as Fig. 11.
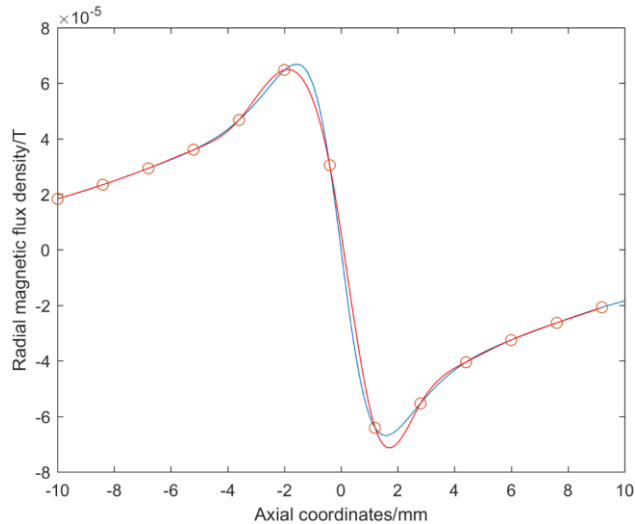
Fig.11. Reconstruction signal

The smoother the signal, the lower the sampling rate required. Additionally, the greater the fluctuation of the signal, the higher the sampling rate required. It can be seen from the previous section that the greater the weld reinforcement, the more intense the permeability change,

the deeper the crack, the lower the liftoff, and the more intense the fluctuation of the signal waveform. When practical factors were considered, the signal sampling rate was maximum when the weld reinforcement was 2 mm, permeability distribution was 4000-3000-2000, crack depth was 4 mm, and liftoff was 3 mm. For the MOI method applied in this study, the detection result is the absolute value of the radial magnetic field. The signal waveform and frequency domain are depicted in Fig. 12.

(a)   (b)

(c)   (d)
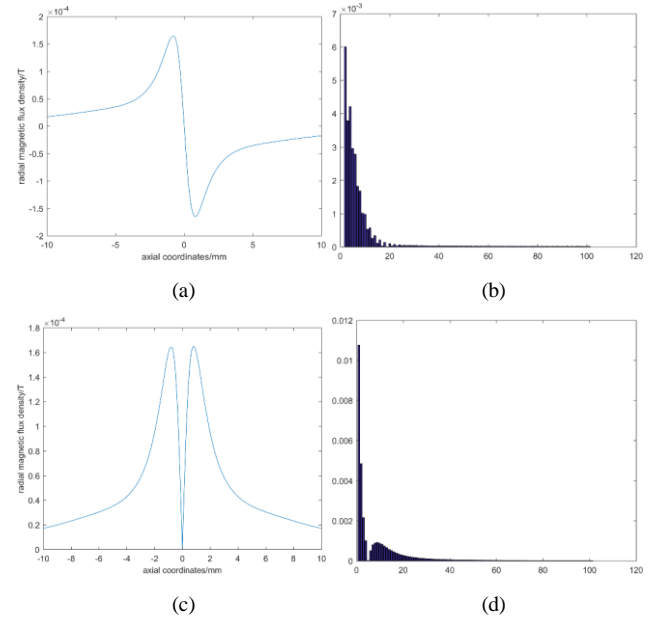
Fig.12. (a) Signal (b) Frequency domain (c) Absolute value signal (d) Frequency domain of absolute value signal

The low-pass cutoff frequency is 0.9 Hz, and the lowest sampling frequency is 1.8 Hz. The lowest spatial resolution is 1.8 times per mm, and the maximum distance between two detection points is 0.55 mm. The reconstruction signal with the 2 Hz sampling rate is shown in Fig. 13.
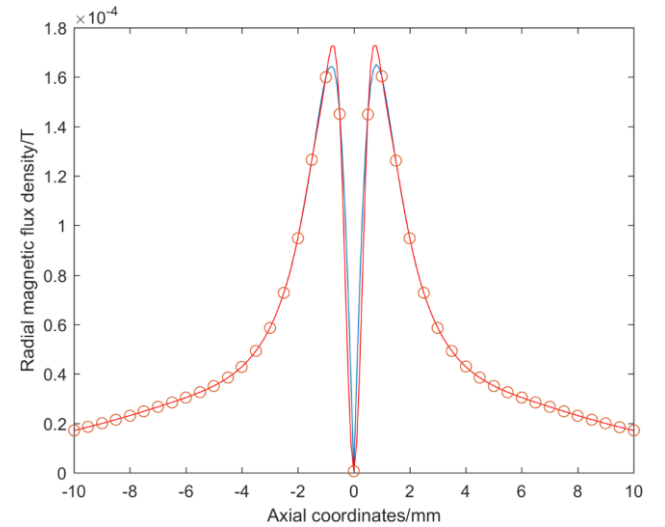
Fig.13. Reconstruction signal

IV.   EXPERIMENT AND ANALYSIS

The weld crack specimen was subjected to experimental tests using the MOI method. The reinforcement height was 2 mm, the weld surface width was 4 mm, the weld crack depth was 4 mm, and the detection liftoff was 3 mm. The specimen and the magneto-optic image are shown in Fig. 14.
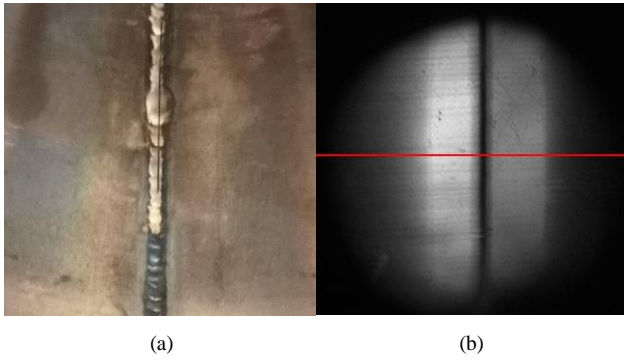
Fig.14. (a) Weld crack (b) Magneto-optic image

The center horizontal line of the magneto-optic image was adopted, and the gray value of this line is shown in Fig. 15(a). This gray value reflects the radial magnetic field intensity of the points. The weld asymmetry resulted in the asymmetry of the signal peaks.
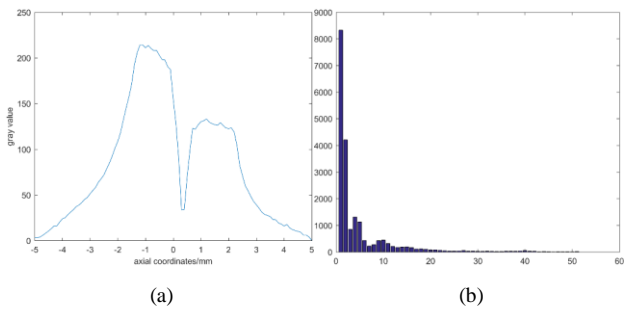
Fig.15. (a) Signal of the weld crack (b) Frequency domain

The cutoff frequency of the low-pass signal was 1.0 Hz, and the lowest sampling rate was 2 Hz. The lowest spatial resolution for reconstructing the signal was two detection points per millimeter, and the maximum distance between adjacent detection points should not exceed 0.5 mm. The experimental results were generally consistent with the simulation results, which proved the preciseness of the finite element simulation.
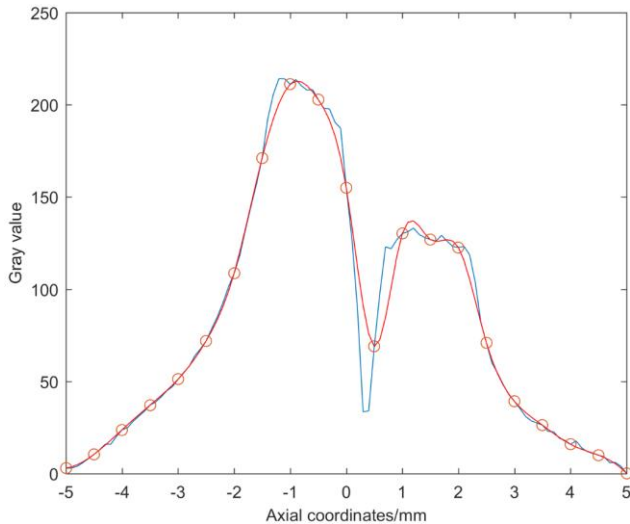
Fig.16. Reconstruction signal

Various interference factors in the experimental tests caused the signal waveform to be more complex than the simulation signal waveform, which required a slightly higher spatial resolution. In practice, the sampling rate should be higher than that of Nyquist to ensure accuracy.

The sampling rate is two times higher the Nyquist frequency, i.e., exceeded 4 Hz. The spatial resolution is more than four points per mm, and the detection point interval is less than 0.25 mm, which maintained that essential information obtained.

V. CONCLUSION

In this study, the frequency-domain analysis method was applied to the MOI method to obtain the lowest spatial resolution while ensuring accuracy. Several significant factors that influence the leakage magnetic field of a weld crack were investigated, and the most complex radial magnetic signal was simulated. Next, the lowest spatial resolution required to ensure the detection accuracy was determined using the most complex signal. This study is useful for reducing the data processing amount of the MOI detection results and provides a reference for the spatial resolution setting of pipeline detectors.

REFERENCES

[1] H. A. Kishawy and H. A. Gabbar, "Review of pipeline integrity management practices", *International Journal of Pressure Vessels and Piping*, vol. 87, no. 7. International Journal of Pressure Vessels and Piping, pp. 373–380, 2010.

[2] Q. Feng, R. Li, B. Nie, S. Liu, L. Zhao, and H. Zhang, "Literature Review: Theory and Application of In-Line Inspection Technologies for Oil and Gas Pipeline Girth Weld Defection", Sensors, vol. 17, no. 12. Sensors, p. 50, 2016.

[3] Z. D. Wang, Y. Gu, and Y. S. Wang, "A review of three magnetic NDT technologies", *Journal of Magnetism and Magnetic Materials*, vol. 324, no. 4. Journal of Magnetism and Magnetic Materials, pp. 382–388, 2012.

[4] Y. Shi, C. Zhang, R. Li, M. Cai, and G. Jia, "Theory and Application of Magnetic Flux Leakage Pipeline Detection", *Sensors*, vol. 15, no. 12. Sensors, pp. 31036–31055, 2015.

[5] A. A. Carvalho, J. M. A. Rebello, L. V. S. Sagrilo, C. S. Camerini, and I. V. J. Miranda, "MFL signals and artificial neural networks applied to detection and classification of pipe weld defects", NDT & E International, vol. 39, no. 8. NDT & E International, pp. 661–667, 2006.

[6] H. R. Vanaei, A. Eslami, and A. Egbewande, "A review on pipeline corrosion, in-line inspection (ILI), and corrosion growth rate models", International Journal of Pressure Vessels and Piping, vol. 149. International Journal of Pressure Vessels and Piping, pp. 43–54, 2017.

[7] L. Tian, Y. Cheng, C. Yin, D. Ding, Y. Song, and L. Bai, "Design of the MOI method based on the artificial neural network for crack detection", *Neurocomputing*, vol. 226. Neurocomputing, pp. 80–89, 2017.

[8] X. Gao and Y. Chen, "Detection of micro gap weld using magneto-optical imaging during laser welding", The International Journal of Advanced Manufacturing Technology, vol. 73, no. 1–4. The International Journal of Advanced Manufacturing Technology, pp. 23–33, 2014.

[9] Y. Cheng, Y. Deng, L. Bai, and K. Chen, "Enhanced Laser-Based Magneto-Optic Imaging System for Nondestructive Evaluation Applications", *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 5. IEEE Transactions on Instrumentation and Measurement, pp. 1192–1198, 2013.

[10] Li, Yanfeng , et al. "Weld cracks nondestructive testing based on magneto-optical imaging under alternating magnetic field excitation." Sensors & Actuators A Physical, vol. 285, pp. 289-299, 2019

Application of Tensor Decomposition Methods In Eddy Current Pulsed Thermography Sequences Processing

Yiping Liang
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
lypddup@163.com

Libing Bai*
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
libing.bai@uestc.edu.cn

Jinliang Shao
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
jinliangshao@126.com

Yuhua Cheng
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
yhcheng@uestc.edu.cn

Abstract—**Eddy Current Pulsed Thermography (ECPT) is widely used in Nondestructive Testing (NDT) of metal defects where the defect information is sometimes affected by coil noise and edge noise, therefore, it is necessary to segment the ECPT image sequences to improve the detection effect, that is, segmenting the defect part from the background. At present, the methods widely used in ECPT are mostly based on matrix decomposition theory. In fact, tensor decomposition is a new hotspot in the field of image segmentation and has been widely used in many image segmentation scenes, but it is not a general method in ECPT. This paper analyzes the feasibility of the usage of tensor decomposition in ECPT and designs several experiments on different samples to verify the effects of two popular tensor decomposition algorithms in ECPT. This paper also compares the matrix decomposition methods and the tensor decomposition methods in terms of treatment effect, time cost, detection success rate, etc. Through the experimental results, this paper points out the advantages and disadvantages of tensor decomposition methods in ECPT and analyzes the suitable engineering application scenarios of tensor decomposition in ECPT.**

Keywords—Nondestructive testing, Eddy Current Pulsed Thermography, ECPT Sequence Processing, Tensor Decomposition.

I. Introduction

Eddy Current Pulsed Thermography (ECPT) is an efficient nondestructive testing (NDT) technology for metal material testing because its large detection area and deeper detection range than traditional detection methods, but the weakness is that the detection results were sometimes affected by coil noise and edge effect noise, therefore, for the better detection results, it is necessary to eliminate the noise interference and extract defects from a complex background.

Some scholars consider the problem of ECPT defect information extraction as a background-modeling problem in the field of image segmentation, which aims to separate the foreground objects from the background. While the large background can be modeled as a low rank element because the gentle temperature change during the eddy current heating process which lead to a high degree of correlation between each frame. On this basis, many sparse matrix decomposition methods have been applied in ECPT. For example, GAO proposed a novel unsupervised sparse component extraction algorithm using the Variation Bayesian framework (VBSA) and verified its effect on steel structural samples [1]. Ahmed combined the sparse dictionary matrix decomposition with the wavelet transform and applied it to the detection of carbon fiber samples [2].

Recently, sparse tensor decomposition is a popular method in the field of image segmentation [3]. Compared with the sparse matrix model, sparse tensor decomposition can analyze the three-dimensional data directly, therefore, the method is suitable for the processing of image sequences (or videos) [4]. For example, Huang used the Sum of Nuclear Norm (SNN) model [5] for sparse tensor decomposition [6] and achieved good effects. Liu proposed the definition of the tensor trace norm for the first time, and then present a completion algorithm based on tensor decomposition to estimate missing values in tensors of visual data [7]. Lu proposed a method of tensor robust principal component analysis (TRPCA) based on a new defined tensor nuclear norm to improve the decomposition effect [8].

Although tensor decomposition has been used in many image segmentation fields, it is rare to find a paper that applies tensor decomposition to the processing of ECPT image sequences at present, among them, only a few literatures compare the tensor decomposition with the matrix decomposition specifically. Therefore, in this paper, the difference between sparse tensor decomposition and sparse matrix decomposition in ECPT is compared in detail by richer and sufficient experiments.

The structure of this paper is as following. The first part introduces the Eddy Current Pulsed Thermography (ECPT), describes the research status of ECPT data processing

---

methods, and summarizes the work in this paper. The second part introduces the theory of sparse tensor decomposition and demonstrates the rationality of the application of it in ECPT. The third part introduces the experimental platform and the samples used in the experiment. The fourth part shows the result of experiments analyzes the results and compared different kinds of methods. The last section concludes the advantages and disadvantages of the application tensor decomposition in ECPT and forecasts the work in future.

II. TENSORS THEORY

A. Tensor's Representation

Tensors are multi-dimensional matrices, only third-order tensors are discussed in this paper. For a third-order tensor A, mode-1,mode-2,mode-3 are often used to express the 'slices' in different directions (see Fig. 1(a)) denoted as A(i,:,:), A(:,i,:), A(:,:,i). Usually, the front slice is used more in decomposition, so it also can be denoted as $A^{(i)}$. A third order tensor can also be treated of a combination of 'fibers'. As shown in Fig. 1(b), 'fibers' have three different directions, which are described by mode-1, mode-2 and mode-3 respectively. In this paper, only mode-3 fibers are used which are denoted as A(i, j,:).
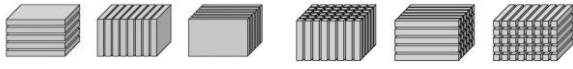
(a) Slices of a third-order tensor   (b) Fibers of a third-order tensor

Fig. 1.Slices and fibers of a third-order tensor

B. Tensor-SVD

Singular Value Decomposition (SVD) is an important matrix decomposition in linear algebra and has been widely used in engineering, especially in principal component analysis. Suppose $A \in \mathbb{R}^{M \times N}$, then SVD have the equation as follows

$$A = USV^* \qquad (1)$$

where U and V are unit orthogonal matrices, S is a diagonal matrix.

Tensor-SVD is an expand of matrix SVD, suppose $L \in \mathbb{R}^{M \times N \times P}$. Applying Fast Fourier Transform to each mode-3 fiber of length P, and the new tensor is denoted as $\bar{L}$. It's easy to see that $\bar{L}^{(i)} \in \mathbb{R}^{M \times N}$. Performing SVD operation (singular value decomposition) for each of the matrices, the results are shown as following

$$\bar{L}^{(i)} = \bar{U}\bar{S}\bar{V} \qquad (2)$$

Take the calculated results U, S and V as the front slice of the tensor and three third-order tensors U, S and V will be obtained. Carry out the Inverse FFT transformation, then the Tensor-SVD results are denoted as

$$L = USV \qquad (3)$$

C. Tensor Robust Principal Component Analysis

Robust principal component analysis (RPCA) is a classical matrix decomposition algorithm witch was proposed firstly by Candes[9], it aims to separate low-rank data from the original data which doped with sparse noise which has been widely used in the field of image processing. In the image matrix, the significant information of the source image can be described by the features extracted from the sparse matrix. In infrared image, the target is the prominent information and the prominent part which is different from the fuzzy background. Therefore, infrared target information can be modeled as the components related to sparse matrix, and background information can be modeled as the components related to low-rank matrix.

The formulaic meaning of RPCA can be described as

$$\min_{L,E} \|L\|_* + \lambda \|E\|_1 , s.t. X = L + E \qquad (4)$$

where L represents a low-rank matrix and E represents a sparse matrix, $\|L\|_*$ represents the nuclear form of the matrix, $\|E\|_1$ represents the l1-norm of the matrix.

Tensor-RPCA(TRPCA) is similar to RPCA, it replaces the matrix with a tensor and trying to decompose it into the sum of two tensors, and the sum of tensors is similar to the sum of matrixes, the sum of the elements at the corresponding positions. In essence, TRPCA aims to find the minimum value of the Eq. (5) which satisfies the following conditions:

$$\min_{L,\mathcal{E}} \|\mathcal{L}\|_* + \lambda \|\mathcal{E}\|_1 \qquad (5)$$

where $\mathcal{L}$ represents the decomposed low-rank tensor, $\mathcal{E}$ represents the decomposed sparse tensor, and $\|\mathcal{L}\|_*$ is the nuclear norm of the low-rank tensor.

TRPCA has different models according to the different definitions of tensor norm. This paper introduces two models in ECPT which are the SNN-model mentioned by Bo Huang in his paper [6] and the TNN-model mentioned by Lu Canyi in his paper [8].

SNN-model uses the sum of nuclear norm as the definition of the tensor norm in Eq. (5) which was first proposed by Liu in his paper [5] as follows

$$\sum_{i=1}^{n} a_i \|L_{(i)}\|_* \qquad (6)$$

where n represents the number of dimensions, in this paper, it was limit to 3, then $L_{(i)}$ means the matrix unfolded by mode-i slice. a_i is constant which satisfying $\sum_{i=1}^{n} a_i = 1$.

In TNN-model, the nuclear norm of tensor L is the sum of diagonal elements of tensor S obtained by Tensor-SVD decomposition. Lu proposed a method[10] to solve the minimum value of Eq. (5). Over the repeated iterative operations, the optimal solution satisfying the optimization function is obtained.

D. The Feasibility Study of the TRPCA's Application in the Field of ECPT

From the Eq. (5), it can be seen that TRPCA decomposes the raw data into a low-rank component and a sparse component. Low rank means that there is a little difference between each slice while sparse is opposite. In the heating stage of ECPT, for a sample with finite uniform thickness, the heat generation rate is defined as the Joule heat generated per unit time, which can be expressed in Cartesian coordinates as follows:

$$Q = \sigma \left[\left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 + \left(\frac{\partial \phi}{\partial z} \right)^2 \right] \quad (7)$$

Here, σ represents potential and φ represents conductivity. Because the two components of the eddy current are expressed by the derivatives of the potential, the heat generation rate Q at the crack tip should theoretically have a trend toward the infinite owing to the discontinuity, and there inevitably form singularities at the crack tip, which lead to a rapid rise in temperature and a high-temperature region [11].

According to the physical laws above, if there is a defect on the surface of a smooth metal, the temperature of the defect area will rise rapidly due to the effect of eddy current heating which will be shown in an image that the gray scale of the pixel will rise rapidly (the whole becomes brighter). Yet, the normal area is less affected by the thermal effect of eddy current, and the temperature rises much slower than the defect area. In ECPT image sequence, this means the gray scale change process of this point is relatively flat.

Therefore, after collecting numerous eddy current thermal images, the pixels of the defects are the prominent information which are obviously different from the overall background, while the pixels of the normal area are the background information with the characteristics of large area and similar imaging features. Using the tensor theory described above, the set of image frames can be described as a third-order tensor, and the defect area is the sparse tensor component contained therein, while the non-defect area is the low-rank tensor component contained therein. Thus the theoretical conditions in TRPCA are highly consistent with the actual scene of eddy current detection. Therefore, TRPCA is suitable to be applied in ECPT sequence processing.

III. EXPERIMENT SETUP

A. The Settings of Experimental Equipments

In order to verify the stability of the methods, the experiments are carried out on two sets of experimental equipments shown in Fig. 2(a) and Fig. 2(b). The difference between the two experimental platforms is the model of IR camera(FLIR SC7500 in Fig. 2(a) and FLIR A655sc in Fig. 2(b)). Specifically, the sampling frequency and maximum output resolution of the two IR cameras are different. FLIR SC7500 supports a resolution up to 240 × 320 and a sampling rate up to 383Hz, while FLIR A655sc supports a resolution up to 480 × 640 and a sampling rate up to 200Hz. Therefore, for different samples, the experiments are carried out on different platforms according to the sample's size and other relevant characteristics (The details are given in the introduction of the samples below).

(a) Experimental setup 1   (b) Experimental setup 2
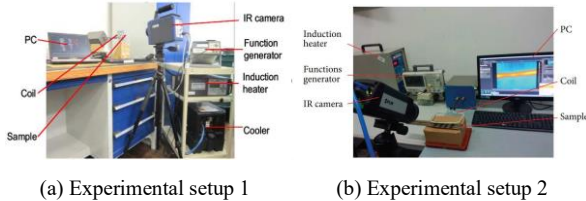
Fig. 2.Experimental setup

Other than that, the configurations of other devices are the same in two sets: A coil is placed directly above the sample and an induction heater passes an alternating current into the coil to generate eddy currents on the surface and inside of the sample. The IR camera is used to record the ECPT sequence on the sample's surface and transmits it to PC for processing. In order to eliminate the influence of the coil's thermal radiation on the experimental results, the coil is bent from a hollow copper tube which forms a watercooled circulation system together with the induction heater. In addition, the outside of the coil is wrapped in a rubber skin so that the coil temperature changes little during the experiment. The data is processed on a PC with 32GB memory and an Intel Quad-Core i5-7500@3.4GHz CPU.

B. Experiment Samples

To fully simulate the various defects, the test samples include four types: A rectangular through-hole sample was used to simulate the surface crack of stainless steel, a tiny round-hole sample was used to simulate the pitting of stainless steel, a multi-defects sample was used to simulate the weld crack of stainless steel, an actual part of the engine blade was used to simulate the thermal fatigue crack. Among them, rectangular through-hole sample and engine blade were tested on the set of Fig. 2(a), tiny round-hole sample and multidefects sample were tested on the set of Fig. 2(b).

1)   surface crack: The sample used to simulate surface cracks is shown in Fig. 3, which is a stainless steel plate with a long and narrow hole through it. The size of sample is 100×45mm2 and the thickness is 0.24mm. The crack is located in the center with a length of 10mm and a width of 2mm. The experiment lasted for 2 seconds and the heating phase lasted 0.1 second. The sampling rate of IR camera was set as 383 Hz, so 766 frames of thermal image video were collected in total. The size of each frame is 256 × 320 pixels. The first 39 frames are the heating process, and the last 726 frames are the heat release process. The first 80 frames were used to construct the tensor of 256 × 320 × 80 for decomposition.

2)   pitting: The sample used to simulate pitting is shown in Fig. 4. The experiments were carried out on defects No.4 (3*mm* radius through hole) and No.7 (5*mm* radius non-through hole) respectively. The experiment lasted for 1.6 seconds with a sampling rate at 200 Hz. The size of each frame is 240×640 pixels. For defect No.4, the 79th frame was corresponded to

Fig. 3.Single rectangular through-hole Sample

the moment when the heating process ended. The frames from 39th to 119th were used to construct a tensor of 240 × 640 × 80 for decomposition. For defect No.7, the 88th frame was corresponded to the moment when the heating process ended. The frames from 48 to 128 were used to construct a tensor of 240 × 640 × 80 for decomposition.

Fig. 4.Tiny round-hole Sample

3) weld crack: The sample used to simulate multi weld cracks is shown in Fig. 5 with six artificial defects. Considering the size of the heating coil and the collection range of the thermograph, this paper collected the heat information of the middle part of the sample including No. 3, No. 4 and No. 5 defects with size of $5 \times 2 \times 5mm^3$, $5 \times 1 \times 1mm^3$ and $5 \times 1 \times 3mm^3$. The collection frequency was 25 Hz, 250 frames were collected and the end time of heating corresponded to 88th frame. The size of each frame was 480×640 pixels. The frames from 48th to 128th were used to construct a tensor of $480 \times 640 \times 80$ for decomposition.
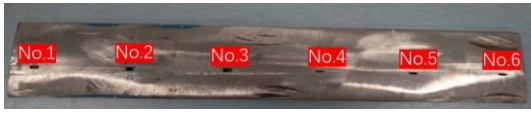

Fig. 5.Multi-defects Sample

4) thermal fatigue crack: The real sample is a turbine engine blade shown in Fig. 6. The experiment lasted for 2 seconds and the heating phase lasted 0.2 second. The sampling rate of IR camera was set as 353 Hz and 766 frames of ECPT sequences were collected in total. 69th frame was the end moment of heating process. The frames from 29 to 109 were used to construct the tensor of $256 \times 320 \times 80$ for decomposition.

C. Quantitative Evaluating Criteria

F-Score is a criteria often used in NDT quantitative evaluation [1], [12], [13] which is calculated by True Positive (TP), False Negative (FN) and False Positive (FP). TP means the defective part indicated by the test results is indeed the defect on the sample. FN means the area where the test result

Fig. 6.Thermal fatigue crack of a steam turbine blade

indicates defect-free but exists defect on the sample. FP means the defective part indicated by the test results is defect-free on the Sample. F-score is defined as follows:

$$F\text{-}Score = \frac{\lambda + 1}{\lambda^2} \times \frac{P \times R}{P + R} \qquad (8)$$

where P and R are defined as

$$P = \frac{TP}{TP + FN}, R = \frac{TP}{TP + FP}. \qquad (9)$$

λ is the weight factor between P and R. The default setting is 1, which means that P and R are equally important.

Mean Square Error (MSE) is also commonly used image quality evaluation parameter which measures the image quality according to the statistical results of pixel error between the image to be evaluated and the reference image. The smaller the MSE value, the better the image quality. MSE is calculated as follows

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} |R(i, j) - F(i, j)|^2 \qquad (10)$$

where F is the image to be evaluated, and R is the reference image, with sizes of $M \times N$.

This paper used F-Score, MSE and time costs for quantitative evaluation.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. Experiments Results

In this paper, EJSLRMD[13] and WIASDMD[2] are used as the representative methods of sparse matrix decomposition applied in ECPT. For the sparse tensor decomposition is seldom used in ECPT, the TNN-TRPCA and SNN-TRPCA mentioned above are used as representatives in this paper. In fact, these two methods are relatively new in the field of image segmentation. All samples are tested respectively, and the results are shown in TABLE I.

Row 2 and Row 3 of TABLE I show the experimental results of EJSLRMD[13] and WIASDMD[2]. It can be seen that these two methods enhanced the display of defects in the original image to a certain extent. In Sample 4 (multi artificial defects), EJSLRMD failed to find all defects. Therefore, this method was not suitable for the analysis of multi-defect specimens. In addition, these two methods were not enough to suppress the pixel values of defect-free area, which can be concluded by comparing the results of Sample 1.

Row 4 and Row 5 of TABLE I shows the experimental results of the SNN-TRPCA and TNN-TRPCA. Comparing with the original images (Row 1 of TABLE I), TRPCA has a good separation effect on the four common defects. For the sample with multiple defects, all the defective parts (including the subsurface defect) can be successfully separated. Most of the noise was suppressed effectively, but some external interference had not been eliminated. Therefore, in addition to the defective parts, some parts without defects were misjudged as defective parts. However, this phenomenon only existed at the boundary of the image, only the edge of the sample was affected. In general, TRPCA has the best effect without considering the time cost and TNN is a little better than SNN.

B. Parameterized Results

In TABLE I, the detection effects of various methods can be roughly compared. However, due to the lack of quantitative evaluation criteria, the evaluation is subjective to a certain extent. In addition, the time costs of methods are not included in TABLE I. To make a more convincing comparison, this section calculates a variety of quantitative indicators and provides a comprehensive comparison of all methods and the results are shown below.

TABLE II shows the time cost of the four methods. It can be intuitively seen that SNN-TRPCA and TNN-TRPCA take much longer time than WIASDMD and EJSLRMD. Among them, TNN performs better than SNN, but in general, the time of tensor decomposition is much higher than that of matrix decomposition.

the extraction effect of defect information. TABLE IV shows the F-Score of the four methods. F-Score is another important quantitative indicator of the detection effect, it combines miss and error detection and reflects the success rate of the algorithm in detecting the defective part. The larger the F-Score, more information of the defect part contained in

TABLE I.   EXPERIMENTAL RESULTS OF ALL SAMPLES. SAMPLE 1 IS THE ARTIFICIAL WELD CRACK SAMPLE SHOWN IN FIG. 3. SAMPLE 2 IS THE ARTIFICIAL THROUGH-PITTING SAMPLE (NO. 4 DEFECT SHOWN IN FIG. 4). SAMPLE 3 IS THE ARTIFICIAL SURFACE-PITTING SAMPLE (NO.7 DEFECT SHOWN IN FIG. 4). SAMPLE 4 IS THE MULTI ARTIFICIAL DEFECTS SAMPLE (NO. 3, NO. 4, NO. 5 DEFECTS SHOWN IN FIG. 5). SAMPLE 5 IS THE NATURAL THERMAL FATIGUE CRACK SAMPLE SHOWN IN FIG. 6.
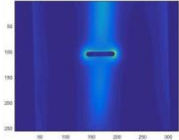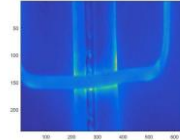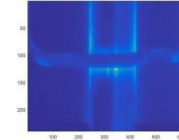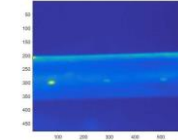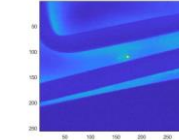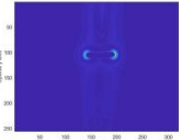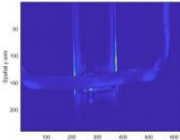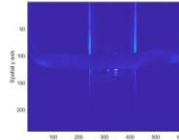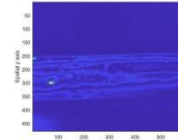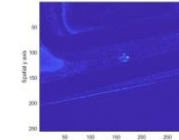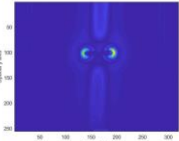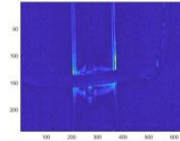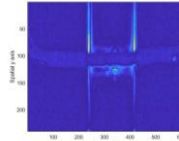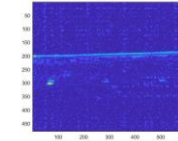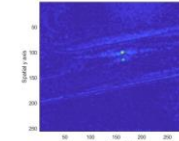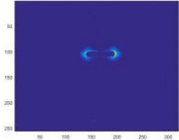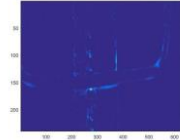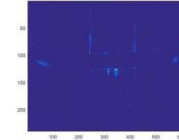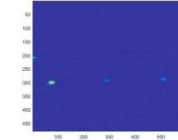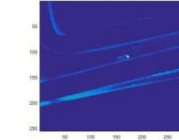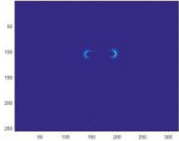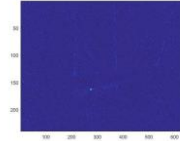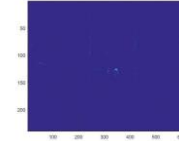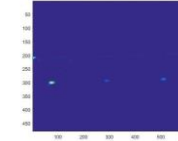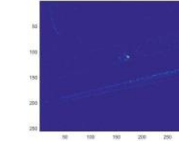
Method	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Original(Row1)					
EJSLRMD(Row2)					
WIASDMD(Row3)					
SNN-TRPCA(Row4)					
TNN-TRPCA(Row5)					

TABLE III shows the MSE of each method. TNN-RPCA has the largest MSE, SNN-TRPCA was following. EJSLRMD[13] and WIASDMD[2] are the smallest. According to the definition of MSE, it can be seen that the smaller the MSE is, the less difference between the separation result and the original image, which means more information from the original image contained in the separation result. The purpose of the above algorithm is extracting defect information from the original image and it is well known that the proportion of defect information in the original image is small. Therefore, when MSE is used to evaluate the detection effect, the larger MSE means the stronger inhibitory effect on non-defective parts. Therefore, TNN and SNN are better than the other two matrix decomposition algorithms in terms of

TABLE II.   RUNTIME COMPARISON (S)

Method	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
EJSLRMD	10.6	19.9	19.7	41.4	10.5
WIASDMD	12.26	21.29	21.3	39.4	11.8
SNN-TRPCA	1196.3	1530.5	1972.5	1557.2	1370.5
TNN-TRPCA	306.9	1196.3	1215.4	985.7	485.7

results. From TABLE IV , it can be seen that SNN-TRPCA is same with TNN-TRPCA and higher than EJSLRMD and WIASDMD.

TABLE III.   MEAN SQUARE ERROR OF METHODS IN SIMULATIONS

Method	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
EJSLRMD	263.9	623.1	215.6	524.3	734.2
WIASDMD	265.5	663.0	172.8	520.5	1050.4
SNN-TRPCA	325.0	917.6	303.2	668.7	1329.6
TNN-TRPCA	356.7	927.1	330.1	657.8	1369.7

TABLE IV.   F-SCORE OF METHODS IN SIMULATIONS

Method	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
EJSLRMD	1	0.5	0.4	0.4	1
WIASDMD	1	0.5	0.5	0.71	1
SNN-TRPCA	1	0.67	1	0.86	1
TNN-TRPCA	1	0.67	1	0.86	1

Besides, it is worth mentioning that both EJSLRMD and WIASDMD contain a manual filtering section. For these two methods, different selections on some parameters will result

in different outputs, so it is necessary to artificially select the appropriate one for different data. To be precise, the results of EJSLRMD are highly affected by the selection of spatialtransient multi-layer and the results of WIASDMD depend on the wavelet level a lot. In this paper, in order to obtain the optimal effect, the layer in EJSLRMD is chosen as 4, 3, 3, 5 and 4 respectively for Sample 1–5, while the wavelet level in WIASDMD is chosen as 2, 3, 3, 2 and 4. Different from the EJSLRMD and WIASDMD, TRPCA produces only a certain output for a certain input. This means that the test results are unique, which is very important in practical industrial applications.

To sum up, comparing with the two methods of matrix decomposition, TRPCA has the advantages of stronger ability to separate defects and higher detection success rate which are reflected in MSE and F-Score respectively. Among the two types of TRPCA, TNN was more effective than SNN. The disadvantages of TRPCA are also obvious, mainly reflected in the time cost, which is often dozens of times of the two matrix decomposition methods. In addition, TRPCA can give unique results for a set of data, so this method eliminate the process of manual screening. From all above, it can be seen that TRPCA is more suitable for automatic detection environment with low speed requirements, high precision and high success rate.

V. Conclusion

This paper demonstrates the feasibility of sparse tensor decomposition theory in the processing of ECPT image sequence and tests TRPCA(a kind of popular tensor decomposition algorithm)'s effect through several experiments on a variety of defects samples, including surface crack, weld crack, pitting and the thermal fatigue crack. Through the comparison with the sparse matrix decomposition method, which has been widely used in ECPT field, this paper, concludes that TRPCA is a high accuracy defect extraction algorithm suitable for ECPT. However, it has a drawback of high time cost, which limits its usage in many situations in need of high processing speed.

Although TRPCA cannot represent all tensor decomposition algorithms, at least it is proved that tensor decomposition is suitable to be introduced into ECPT. In the future, we will carry out more works on the combination between tensor decomposition and ECPT to get better results.

References

[1] B. Gao, P. Lu, W. L. Woo, G. Y. Tian, Y. Zhu, and M. Johnston, Variational Bayesian Subgroup Adaptive Sparse Component Extraction for Diagnostic Imaging System, IEEE Trans. Ind. Electron., vol. 65, no. 10, pp. 8142–8152, 2018.

[2] J. Ahmed, B. Gao, and W. L. Woo, Wavelet-Integrated Alternating Sparse Dictionary Matrix Decomposition in Thermal Imaging CFRP Defect Detection, IEEE Trans. Ind. Informatics, vol. 15, no. 7, pp. 4033–4043, Jul. 2019.

[3] T. G. Kolda and B. W. Bader, Tensor Decompositions and Applications, SIAM Rev., vol. 51, no. 3, pp. 455–500, Aug. 2009.

[4] A. Cichocki et al., Tensor Decompositions for Signal Processing Applications: From two-way to multiway component analysis, *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.

[5] J. Liu, P. Musialski, P. Wonka, and J. Ye, Tensor Completion for Estimating Missing Values in Visual Data, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[6] B. Huang, C. Mu, D. Goldfarb, and J. Wright, Provable models for robust low-rank tensor completion, *Pacific J. Optim.*, vol. 11, no. 2, pp. 339–364, 2015.

[7] J. Liu, P. Musialski, P. Wonka, and J. Ye, Tensor Completion for Estimating Missing Values in Visual Data, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[8] C. Lu, J. Feng, S. Yan, and Z. Lin, A Unified Alternating Direction Method of Multipliers by Majorization Minimization, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 3, pp. 527–541, Mar. 2018.

[9] E. J. Cands, X. Li, Y. Ma, and J. Wright, Robust principal component analysis, *J. ACM*, vol. 58, no. 3, pp. 1–37, May 2011.

[10] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm, *IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 4, pp. 925–938, Apr. 2020.*

[11] T. Sakagami and S. Kubo, Development of a new crack identification method based on singular current field using differential thermography, *Thermosense XXI*, 1999, pp. 369–376.

[12] B. Gao, X. Li, W. L. Woo, and G. Y. Tian, Physics-Based Image Segmentation Using First Order Statistical Properties and Genetic Algorithm for Inductive Thermography Imaging, *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2160–2175, May 2018.

[13] J. Ahmedtl, B. Gao, W. L. Woo, and Y. Zhu, Ensemble Joint Sparse Low Rank Matrix Decomposition for Thermography Diagnosis System, *IEEE Trans. Ind. Electron.*, to be published, doi: 10.1109/tie.2020.2975484

A Portable NeuECG Monitoring System for Cardiac Sympathetic Nerve Activity Assessment

Yantao Xing
*School of Instrument Science and
Engineering
Southeast University*
Nanjing, China
230198304@seu.edu.cn

Jianqing Li*
*School of Instrument Science and
Engineering
Southeast University*
Nanjing, China
ljq@seu.edu.cn

Zhenyuan Hu
*School of Instrument Science and
Engineering
Southeast University*
Nanjing, China
220173289@seu.edu.cn

Yuwen Li
*School of Instrument Science and
Engineering
Southeast University*
Nanjing, China
liyuwen@seu.edu.cn

Yike Zhang
*Division of Cardiology
The First Affiliated Hospital of Nanjing
Medical University*
Nanjing, China
zhangyike@njmu.edu.cn

Chang Cui
*Division of Cardiology
The First Affiliated Hospital of Nanjing
Medical University*
Nanjing, China
cuichang@njmu.edu.cn

Cheng Cai
*Division of Cardiology
The First Affiliated Hospital of Nanjing Medical University*
Nanjing, China
15651638971@163.com

Chengyu Liu*
*School of Instrument Science and Engineering
Southeast University*
Nanjing, China
chengyu@seu.edu.cn

Abstract—**Cardiac sympathetic nerve activity is closely related to arrhythmia. Simultaneous non-invasive recording of electrocardiogram and skin sympathetic nerve activity (neuECG) is a new method for sympathetic nervous system assessment. This study presented a data acquisition and analysis system based on neuECG for real-time assessment of cardiac sympathetic nerve activity. Owing to weak skin sympathetic nerve activity (SKNA) signals, signal-acquisition equipment is designed, which is characteristics of low-noise, high-precision and portable. Through a series of experiment, the practicality and reliability of this system are verified. Results show that there are good linear relationships between SKNA signals collected by the proposed system and the normal system (PowerLab), the correlation coefficient is 0.96. Compared with the main neuECG monitoring systems, the proposed system has advantages such as smaller volume and much lower background noise. The system realized the effective and portable evaluation of the cardiac sympathetic activity, in order to identify the risk of arrhythmia as soon as possible.**

Keywords—Measurement, Internet of Things, Bluetooth, Monitoring, neuECG

I. INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of death globally (32.1%), 17.9 million people have died because of CVDs in 2015 [1]. However, 90% of CVDs is preventable if the disease is detected as soon as possible [2]. CVDs usually lead to autonomic neuropathy, manifested as disorder or structural damage of sympathetic and vagal innervation. Since the autonomic nervous system takes part in various physiological processes, it is necessary to assess autonomic nervous activity for the disease diagnosis and risk stratification for patients .

Traditional studies of estimating sympathetic nerve such as microneurography [5] and heart rate variability (HRV) [3] are either invasive [6] or require proper sinus node function [4]. Thus, these methods are difficult to perform effective and long-term sympathetic nerve monitoring and evaluation. A research team from Indiana University [7] recently proposed a new method named as neuECG [14], i.e., the simultaneous non-invasive recording of ECG and skin sympathetic nerve activity (SKNA) [8]- [10], which can record sympathetic nerve activity directly. Sympathetic activity can be quantified by analyzing different filter frequencies (0.5-150 Hz cutoff frequency for ECG, 500-1000 Hz cutoff frequency for SKNA) [11]. Based on this new method, SKNA bursts was demonstrated that it is critical for indicating the initiation and termination of cardiac arrhythmias.

However, based on lacking dedicated monitoring system [13]，neuECG has limitation in practical application [14], such as sympathetic nerve assessment before and after atrial fibrillation ablation, and daily sympathetic nerve monitoring of patients. Therefore, a portable neuECG real-time assessment system has important significance in the treatment and prevention of patients' diseases. In addition, neuECG can also be extended to sleep quality assessment [11], rehabilitation therapy for patients with Parkinson's disease [12].

In order to develop this system, there are still some problems need to be resolved out. The first problem is that neuECG requires high signal quality. Impulses produced by sympathetic nerve are extremely weak (about 1 μV), and the amplitude of SKNA is around ± 60 μV, which is easy to be contaminated by noise. Thus, reducing the noise interference of the signal and obtaining the signal accurately is challenging in the current research. Meanwhile, due to the miniaturization [15] and low-power design [17], the circuit board layout is compact, so it is more susceptible to noise interference from the environment and electrical equipment [16]. In summary, it is one of our challenges in achieving miniaturization of equipment and high-quality signal acquisition.

The second comes from the algorithm of neuECG analysis and processing. SKNA can reflect sympathetic nerve activity indirectly, the raw SKNA signal can only provide an approximate shape of neural activity. Therefore, quantifying

the changes and morphology of the SKNA outbreak and linking the results to physiological events is the difficulty of current research.

In this study, first we described the system architecture of the developed neuECG assessment system. Then we focused on the experimental verification of sympathetic nerve assessment method based on this portable system. We aimed to present a long-term and real-time sympathetic nerve monitoring scheme with integrated circuit and Internet of Things (IoT) technology [18] - [21], [25].

This paper is organized as follows. The system architecture for the developed sympathetic nervous Estimating system was summarized in section II. Section III presents the experiment designs. Section IV details the experimental results. Section V gives the discussions and conclusion.

II. ARCHITECTURE OF PORTABLE NEUECG SYSTEM

In this section, we will introduce a neuECG monitoring system which provided 24-hour real-time sympathetic nerve activity assessment [22]. The system consists of a miniaturized device that records signal in real time, a gateway, and a software platform based on sympathetic nerve monitoring. The small device worn by the human body is the core component of the system. Miniaturization, low cost and clean signal are the characteristics of this device. It can record simultaneously ECG and SKNA, and transmit the acquired data by wire or wirelessly to data analysis software for display and further processing.
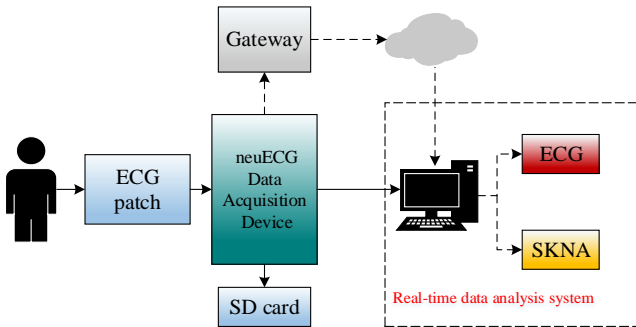
Fig. 1. Block diagram of the neuECG Data Acquisition and Analysis System for Non-invasive Skin Sympathetic Nerve Activity Assessment.

Use the traditional silver/silver-chloride (Ag/AgCl) electrodes to paste to the corresponding location, and the raw data can be recorded by TF card or sent to the PC for display by USB port. The data can also be transmitted to the server through the gateway. Professionals can obtain the data through the server and use the software platform for data analysis. Fig. 1 shows the block diagram of the neuECG system. This system design provides flexibility for a set of use-case scenarios.

A. System Architecture

The data collection device is the core component of the system, consists of front-end circuit, data transmission, memory, power management, user interface and time stamp. Fig. 2 shows the main elements of the data acquisition system [23]. The data collection is done by a low-noise analog front end (AFE) . The microcontroller unit (MCU) control the data acquisition, storage and transmission. The system has one button to control the system and the other one to record the time stamp. The current status of system was a show on the

OLED display. All system modules are actuated by the power management circuit.
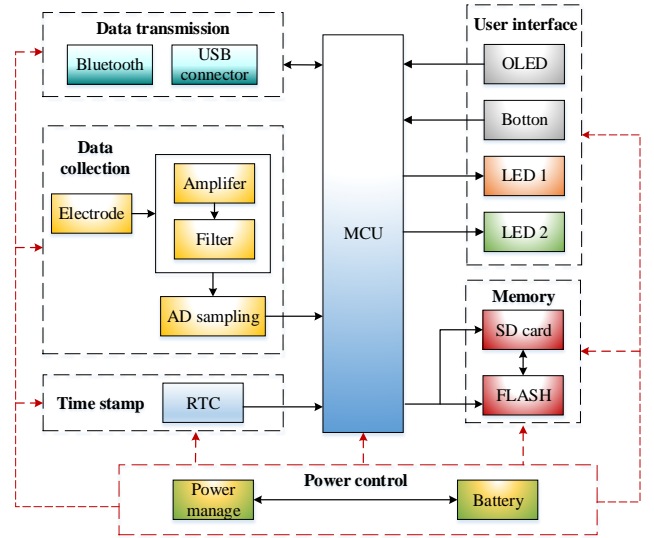
Fig. 2. Block diagram of the data acquisition device for collecting neuECG signals

B. Components Selection and Design Rationale

1) Data Collection Front-End Circuit

The bio-electric potential sensed from the body surface is collected by the traditional Ag/AgCl electrodes. The amplitude of neuECG is of low value, from 0.05 mV up to 4 mV for ECG, from 0.5 µV up to 100 µV for SKNA. Therefore, it requires a high-precision Analog-to-Digital converter to be sampled. So we selected the ADS1299 module of Texas Instruments as the amplifier. This amplifier enables neuECG monitoring with sufficient precision due to its integrated 24-bit ADC.

We chose a STM32F103 microcontroller from STMicroelectronics for the MCU because of its high performance, low consumption and cost. The controller architecture allows to Improve battery life in portable measuring applications with three low-power modes.

2) Peripherals Circuit: Power Management, Data Storage and Transmission

The system uses Bluetooth, USB 2.0 and TF card for data storage. We chose a Bluetooth 2.0 module for wireless transmission with a footprint of 16×11.6 mm. The interaction with the microcontroller is realized through a UART interface. Meanwhile, an TF card storage module based on the FatFS file system is also applied to the system. When the device uses a TF card for data storage or data cable for data transmission, the power consumption is 20mA. When the device transmits data in Bluetooth mode, the power consumption is 50mA. Since the power consumption is high when three storage methods are used at the same time, we designed a switch to freely control the use of these data transfer and storage methods.

All the elements of electrical circuit are powered by a 3.7 V battery which has 3000mAh battery capacity.

3). Noise Suppression Design

To remove power frequency noise and interference, we have optimized as follows.

NeuECG require a bandwidth window of 0.05–1000 Hz. a band pass filter is used for it. Meanwhile, we designed an analog notch filter and a digital notch filter to eliminate possible power frequency interference.

We chose the cable with shielding layer as the lead wire of the system to suppress the interference of environmental noise. In order to suppress random noise and increase the stability of the system, we adopted the PCB design of the multilayer board. Although the manufacturing cost has increased, this has improved the shielding effect of the system. In addition, the design of the multilayer PCB board can further reduce the area of the circuit board.

C. Assembled Prototype for the NeuECG Data Acquisition

The system design meets the requirements with the 0.05–1000 Hz bandwidth, 24-bit ADC resolution and 4/8 kHz sampling frequency. Table 1 summarizes the technical specifications of the system. Over 16 Gb of memory storage, together with 3000mAh battery capacity enable long-time monitoring for more than 2 days. The proposed system records signal in the TF card, or you can choose to send the data to the analysis and processing software through the data line or gateway.

TABLE I TECHNICAL PARAMETER OF THE PROPOSED SYSTEM FOR NEUECG MONITORING.

Parameter	Value
Physiological measurements	ECG, heart rate, SKNA
Band-width	0.05 Hz-1 kHz
Sampling frequency	2 kHz/4 kHz/8 kHz
Number of channels	8
ADC Resolution	24 bit
Memory capacity	16 Gb for TF card
Battery capacity	3000 mAh

Fig.3a shows the assembled PCB of the system, comprising data collection front-end circuit, power management circuit, data storage and transmission module. The developed neuECG data acquisition device is shown in Fig. 3b，with the designed hardware we obtained neuECG signal, displayed on the console.

III. EXPERIMENTAL DESIGN

In order to verify the accuracy and reliability of the device, we prepared two sets of experiments for verification. The data was obtained from human subjects. The human subject study protocols have been approved by the Ethics Committee of The First Affiliated Hospital of Nanjing Medical University.

The research object of the experiment is healthy volunteers and the patient. We record signals with PoweLab Data Acquisition Hardware Device (ADInstruments) and the developed neuECG assessment system simultaneously. The data acquired from PowerLab Data Acquisition Hardware Device (DAQ) was analyzed by LabChart pro 8 software (ADInstruments). The data acquired from the developed system was analyzed by the real-time data analysis system. The data were recorded by traditional Ag/AgCl electrodes, and a conductive adhesive was applied to the skin surface before the experiment began. In the experiment, the recording time, sampling frequency, and band-pass filter settings of the two devices were the same.
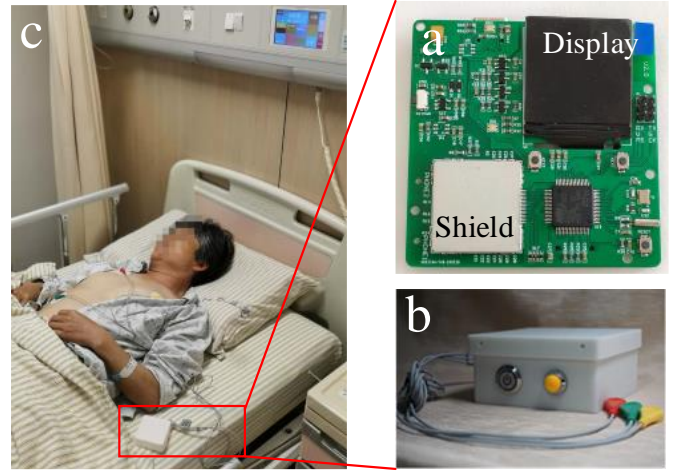
Fig. 3. (a) The assembled PCB of neuECG Acquisition Hardware Device. (b)The Developed neuECG device with Electrode wire. (c) The neuECG assessment system used on a patient's chest in The First Affiliated Hospital of Nanjing Medical University.

The first part of data come from healthy volunteers during the Valsalva Maneuver (VM). VM is a standard procedure for triggering sympathetic discharges. The time of the whole maneuver is generally 20 to 30 s. Therefore, in order to form a positive and negative control of the Valsalva Maneuver, we set the recording time of the experiment to 60s. We simultaneously collect the signals from volunteers with the developed device and PowerLab DAQ .

The second set of signal experiments was performed on patients in the hospital. The patient came from The First Affiliated Hospital of Nanjing Medical University and was unconscious at that time. The experiment mainly recorded the data of patients during sputum suctioning. During sputum suction surgery, the strong stimulation will cause different changes in sympathetic nerves. The purpose is to verify that the system can observe changes in the patient's nerve burst during clinical events clearly.

IV. RESULTS

Using the neuECG data acquisition device, we can clearly see the changes in the SKNA and ECG signal during the operation. Unlike previous complex sympathetic nerve detection systems, the new system is smaller and has more diverse applications. And the low-cost hardware and the design of cloud architecture makes medical mega data possible. The new monitoring system can conveniently enhance the recognition of sympathetic nervous system real-time change. The new system can now be used for a wide range of experiments and surgical monitoring. the reliability and usefulness of the system are verified with the test data.

The experiment was conducted 20 times, and we analyzed it with one of them. Fig. 4 shows ECG and SKNA captured from healthy volunteers using the developed device and PoweLab DAQ simultaneously during the Valsalva Maneuver. Increased SKNA were evident during Valsalva maneuver (VM); The purpose of this experiment is to prove the effectiveness and stability of the system, and the experimental data proves this. The baseline of the ECG collected by the developed system is more stable than PoweLab, which provides a better raw signal for analyzing HRV. Moreover, the SKNA collected by the developed device has lower noise than the noise of PoweLab DAQ. In order to intuitively quantify SKNA and compare the noise baseline, we verify by

calculating eSKNA [13], which can display periods of nerve activation in the given period. The eSKNA was calculated as

Set the input data as a one-dimensional array X.

$$X = [x_0, x_1, \dots, x_n] \qquad (1)$$

set 100 ms as window size, 50 Hz as the moving frequency, and we calculate the average of signal values in the selected window, then, move the window to the next 20 ms, repeats the steps until the end of the data. The equation is:

$$MA = [ma_0, m_{a1}, \dots, ma_j] \qquad (2)$$

$$ma_a = \frac{\sum_{i=a}^{m+a-1} x_i}{m} \qquad (3)$$

$$a = 0, 1, \dots, j \qquad (4)$$

$$j = \frac{(SR \times T - m) \times M_{freq.}}{SR} \qquad (5)$$

where m is the number of samples in a window; j is the number of MA, x_i is the ith sample value of array X; $M_{freq.}$ is moving frequency; SR is sampling frequency of data; T is the duration in second of selected data.

Signal was effectively downsampled 200 times after the operation. Then we pass the MA through root mean square(RMS) calculator. Equations (4), (5) still apply.

$$RMS = [rms_0, rms_1, \dots, rms_j] \qquad (6)$$

$$rms_a = \sqrt{\frac{\sum_{i=a}^{n+a-1} ma_i^2}{n}} \qquad (7)$$

where n is the number of samples in a window; j is the number of samples of RMS, rms_a is the ath sample value of RMS; ma_i is the ith element of moving average array.

After the process, RMS is defined as "eSKNA" which represents the main trend of nerve activity.
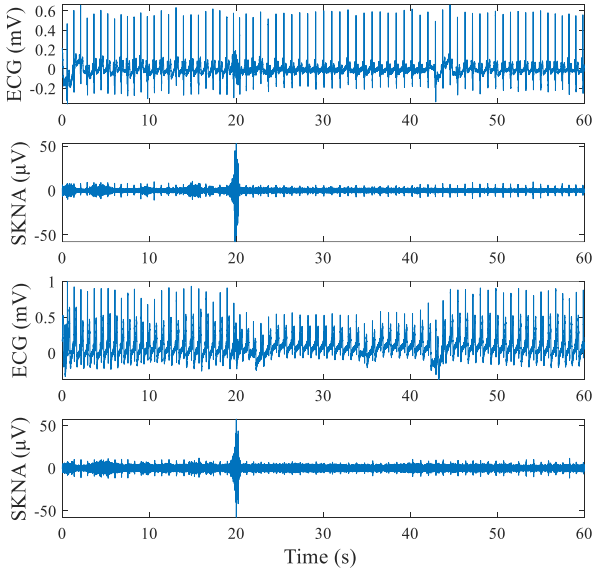
Fig. 4. Signal collected from healthy volunteer with the proposed device and PoweLab DAQ during VM. (a) ECG collected from healthy volunteer with the proposed device during VM. (b) SKNA collected from healthy volunteer with the proposed device during VM. (c) ECG collected from a healthy volunteer with PowerLab DAQ during VM. (d) SKNA collected from a healthy volunteer with PowerLab DAQ during VM.
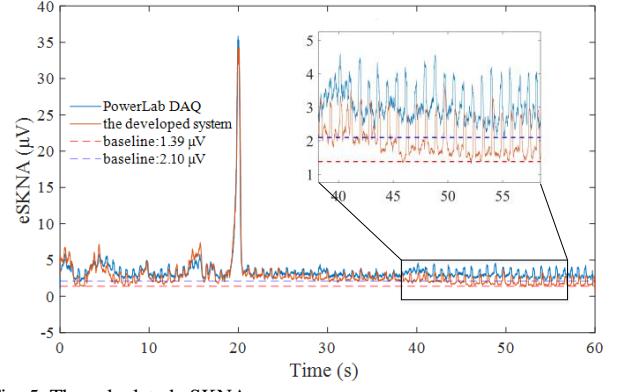
Fig. 5. The calculated eSKNA

The experiment was conducted twenty times in total, and the calculation results are as follows.

TABLE II CALCULATION RESULT OF COMPARISON EXPERIMENT

	PowerLab DAQ Baseline (µV)	NeuECG system Baseline (µV)	Correlation
1	2.10	1.39	0.9541
2	1.86	1.27	0.9712
3	1.79	1.34	0.9773
4	2.03	1.12	0.9526
5	1.89	1.33	0.9468
...	...	...	...
mean	1.89	1.29	0.9655

Fig. 5 shows the calculated eSKNA using the developed device and PoweLab DAQ simultaneously during the Valsalva Maneuver. We used the baseline of eSKNA to compare the noise floor of two devices actually tested. We can see that the baseline came from the developed system is less than PowerLab DAQ's, this shows that the developed system has a better noise suppression level. The trend of signal changes is almost the same. The correlation coefficients as the correlation of the two signals was calculated as

$$r(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var[X]Var[Y]}} \qquad (8)$$

$Cov(X, Y)$ is the covariance of X and Y, $Var[X]$ is the variance of X, $Var[Y]$ is the variance of Y.
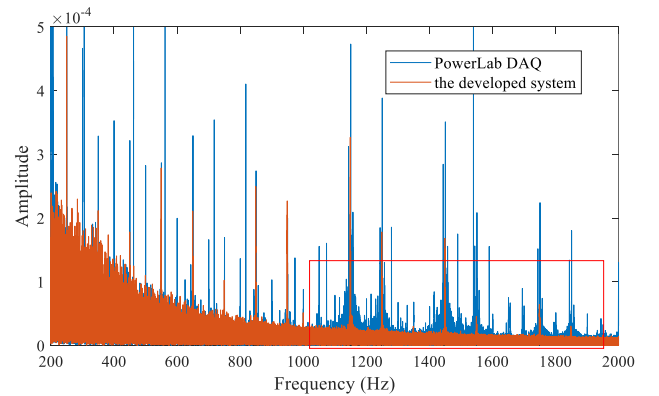
Fig. 6. Spectrogram of the raw signal

Through twenty groups of experiment, the practicality and reliability of this system are verified. According to the results from Table II, there are good linear relationships between SKNA signals collected by the proposed system and the standard system (PowerLab), the correlation coefficient is 0.9655. In addition, by comparing the frequency spectrum of the two raw signals [Fig. 6], we found that the spectrum

energy at 1000-2000 Hz of the signal collected by Powerlab is higher, indicating higher noise. It means the developed system has the lower noise. The results show that this proposed system gets a better effect in comparison with the normal system (PowerLab) of records SKNA non-invasively.

Fig. 7 shows the results from one of the patients during the sputum suctioning. Compared to the preparing and recovery period, the SKNA signal was activated in the operation duration. Dotted red lines mark start and stop of the operation. This result provides an evidence to verify sensitivity of the system. Thus, it can be verified that the neuECG data acquisition and analysis system can be used as the tool to survey SKNA in daily sympathetic nerve monitoring.
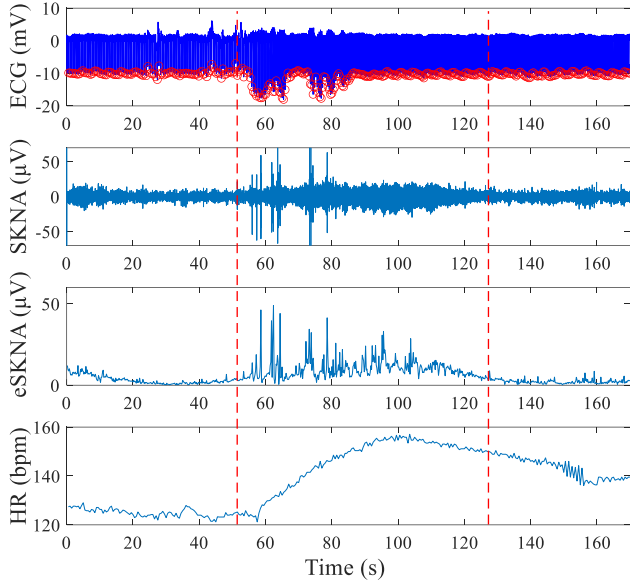
Fig. 7. test data and its system outputs during the suctioning. Red dotted lines mark start and stop of the operation. a: ECG and R peak detection; b: SKNA (µV); c: eSKNA (µV); d: Heart rate / bpm;

V. DISCUSSION & CONCLUSION

The autonomic nervous system plays a very important role in human daily life and lots of diseases are related to it. Traditional autonomic nerve analysis and evaluation methods have limitations. Recently a new method was proposed for the neuECG. However, this method has no special monitoring system, which makes the research of this method difficult to popularize. We summarize the current monitoring system used in this method and demonstrate it in Fig. 8, with the summary of their characteristics compared to ours in Table III. The technical parameters in the table are from the technical manual of the device.

ME6000 Biomonitor [Fig. 8(a)] is the signal acquisition system used at the beginning of neuECG research [8]. This system is mainly used to measure EMG, and it has moderate dimensions and can be used to acquit SKNA. However, its sampling frequency can only be up to 2kHz, and the collected signal integrity is insufficient. In addition, the continuous use time of the system is only 8 hours, so it cannot achieve long-term sympathetic nerve monitoring.

PowerLab DAQ [Fig. 8(b)] is a signal analysis instrument used in neuECG research protocol [14]. It provides high-quality signals and plenty of signal processing algorithms, however, it has poor portability and high cost, and this instrument can only use USB 2.0 to transfer data. So it can

only be used for limited research. Thus, it hardly meets the personalized, multi-scenario monitoring requirements.

(a)

(b)

(c)
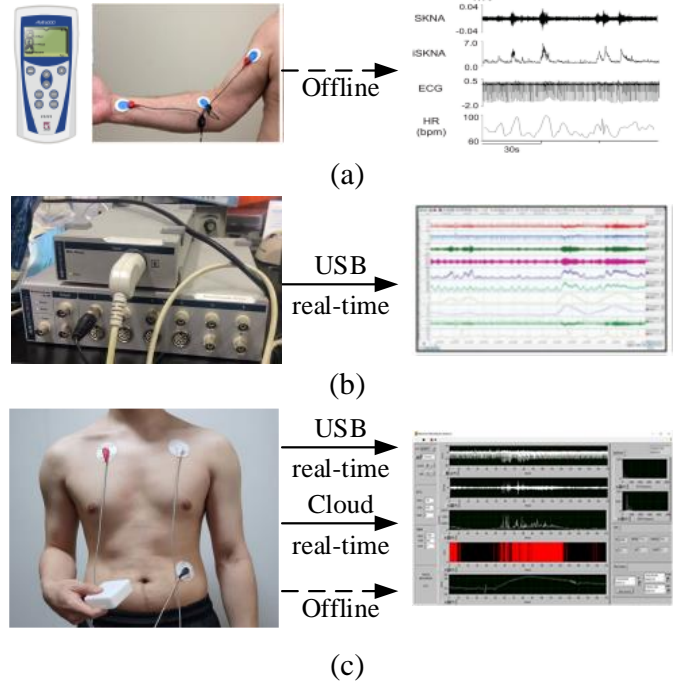
Fig. 8. The current monitoring system used for neuECG. (a) ME6000 Biomonitor. (b) PowerLab Data Acquisition Hardware Device. (c) the development neuECG assessment system.

TABLE III SYSTEM CHARACTERISTICS OF THE MONITORING DEVICE FOR NEUECG RECORDING

Characteristics	ME6000 Biomonitor	PowerLab DAQ	neuECG system
Signal quality	clean	clean	clean
Dimensions	181×85×35 mm	260×240×70 mm	70×70×40 mm
Weight	344 g	2700 g	130 g
Monitoring time	8 hours	AC power supply	> 48 hours
Noise (rms)	1.6 µV	2.2 µV	0.8 µV
Signal real-time analysis?	No	Yes	Yes
Cloud architecture?	No	No	Yes
Portable?	Yes	No	Yes
Cost	Middle	High	Low

In summary, in order to build a big data neuECG monitoring and processing platform, we have developed a portable neuECG monitoring system [Fig. 8(c)]. The advantages of the new portable neuECG monitoring system are: 1) Light weight, compact size, low-cost design and long battery life provide a more flexible solution for sympathetic nerve monitoring. 2) we developed a cloud platform for neuECG monitoring and processing; 3) At last, we focused on the neuECG signal, designed a data acquisition system that can collect high-quality original data, and verified the performance through comparative experiments.

Signals can be observed by this system in real time, and the relationship between Sympathetic nerve and specific episodes of cardiac events can be found. When a patient suffers from an autonomic nervous system-related disease, not only can be monitored in hospital, but also at home or anywhere else. Doctors can observe the patient's sympathetic nerve signal changes in real time in the cloud, thus giving an appropriate treatment plan. This is helpful for disease

prevention and treatment. In addition, the significance of real-time cardiac sympathetic nerve activity assessment is that it can correspond more frequently to clinical events and neuECG signal changes, thereby expanding the field of neuECG research. Thus, the system can provide auxiliary means for research in the field of sympathetic nerve monitoring and has promising application foreground.

REFERENCES

[1] Cardiovascular Diseases (CVDs), World Health Org., Geneva, Switzerland, 2018.

[2] H McGill, C McMahan, "Preventing heart disease in the 21st century: Implications of the pathobiological determinants of atherosclerosis in youth (PDAY) study," *Circulation*, vol. 117, no. 9, pp. 1216-1227, 2008.

[3] M Wolf, G Varigos, et al., "Sinus arrhythmia in acute myocardial infarction," *Aus. Med. J.*, vol. 2, no.2, pp.52-53,1978.

[4] D Linz, F Mahfoud, et al., "Renal sympathetic denervation provides ventricular rate control but does not prevent atrial electrical remodeling during atrial fibrillation," *Hypertension*, vol. 61, pp. 225-231, 2013.

[5] K Hagbarth, A Vallbo, "Pulse and respiratory grouping of sympathetic impulses in human muscle-nerves," *Acta Physiol. Scand.*, vol.74, pp. 96-108, 1968.

[6] B Jung, A Dave, et al., "Circadian variations of stellate ganglion nerve activity in ambulatory dogs," *Heart Rhythm*, vol. 3, pp. 78-85, 2006.

[7] E. A. Robinson, K. S. Rhee, et al., "Estimating sympathetic tone by recording subcutaneous nerve activity in ambulatory dogs," *J. Card. Elec.*, vol. 26, pp. 70-78, 2015.

[8] A Doytchinova, J Hassel, et al., "Simultaneous Non-Invasive Recording of Skin Sympathetic Nerve Activity and Electrocardiogram," *Heart Rhythm*, vol. 14, pp. 25-33, 2017.

[9] M Shen and D Zipes, "Role of the autonomic nervous system in modulating cardiac arrhythmias," *Circ. Res.*, vol. 114, pp. 1004-1021, Mar 14 2014.

[10] S Zhou, B Jung, et al., "Spontaneous stellate ganglion nerve activity and ventricular arrhythmia in a canine model of sudden death," *Heart Rhythm*, vol. 5, pp. 131-9, 2008.

[11] K Miki, and M Yoshimoto, "Differential effects of behaviour on sympathetic outflow during sleep and exercise," *Exp. Physiol.*, vol. 90, pp. 155-158, 2005.

[12] V Arnao, A Cinturino, et al., "In patient's with Parkinson disease, autonomic symptoms are frequent and associated with other non-motor symptoms," *Clin. Auton. Res.*, vol. 25, pp. 301-307, 2015.

[13] C Liu, J Wong, et al., "Method for Detection and Quantification of Non-Invasive Skin Sympathetic Nerve Activity," *2018 International Conference on System Science and Engineering*, 2018.

[14] T Kusayama, J Wong, et al., "Simultaneous noninvasive recording of electrocardiogram and skin sympathetic nerve activity (neuECG)," *Nature Protocols*, vol. 5, no, 15, pp. 1853-1877, 2020.

[15] X Yang, G Liu, et al., "The history, hotspots, and trends of electrocardiogram," *J. Geriatr. Cardiol.*, Vol. 12, no.4, pp. 448-456, 2015.

[16] S Amendola, R Lodato, et al., "RFID technology for IoT-based personal healthcare in smart spaces," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 144-152, 2014.

[17] S Izumi, K Yamashita, et al., "A wearable healthcare system with a 13.7 μA noise tolerant ECG processor," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 5, pp. 733-742, 2015.

[18] J Ding, Y Tang, et al., "A Novel front end design for bioelectrical signal wearable acquisition," *IEEE Sensors J.*, vol. 19, no. 18, pp. 8009–8017, 2019.

[19] C Liu, X Zhang, et al., "Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1363-1374, 2019.

[20] C Liu, M Yang, et al., "Wearable ECG: History, Key Technologies and Future Challenges," *Chinese Journal of Biomedical Engineering*, 2019, 38(6): 641-652.

[21] Y Zhang, L Sun, et al., "Ubiquitous WSN for healthcare: Recent advances and future prospects," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 311-318, Aug. 2014.

[22] N Glazkovaa, T Podladchikovaa, et al., "Non-invasive wearable ECG-patch system for astronauts and patients on Earth," *Acta Astr.*, vol. 166, pp. 613-618, 2020.

[23] B Taji, S Shirmohammadi, et al., "Impact of Skin–Electrode Interface on Electrocardiogram Measurements Using Conductive Textile Electrodes," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 6, pp, 1412–1422, 2014.

[24] A Sikora, M Schappacher, and V Groza, "Evaluation of A Novel Ultra-Low Energy Real-Time SmartMAC Protocol." *in 2019 IEEE International Instrum. Meas. Tech. Conf. (I2MTC)*, 2019.

[25] O Ozhan, Y Karadana, et al., "A Portable Wearable Tele-ECG Monitoring System," *IEEE Trans. on Instrum. Meas.*, vol. 69, no. 1, pp. 173-182, 2020.

An Octave Convolution Neural Network-based QRS Detector

Wei Liu
School of Instrument Science and Engineering
Southeast University
Nanjing, China
213173146@seu.edu.cn

Xingyao Wang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
xingyao@seu.edu.cn

Hongxiang Gao
School of Instrument Science and Engineering
Southeast University
Nanjing, China
gaohx@seu.edu.cn

Chenxi Yang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
chenxiyang@seu.edu.cn

Jianqing Li
School of Instrument Science and Engineering
Southeast University
Nanjing, China
ljq@seu.edu.cn

Chengyu Liu*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
chengyu@seu.edu.cn

Abstract—**Detection of QRS complex in electrocardiogram (ECG) is the most critical and basic step for automated cardiac diagnosis. Many automated QRS complex detection methods have been proposed, and several state-of-art approaches show acceptable detection accuracy. However, current methods cannot generalize to out-of-distribution data, especially inter-patient and low-quality wearable ECGs.**

In our work, a 20-layer convolutional neural network (CNN) using octave convolution was proposed to detect QRS complex in unknown and noisy ECGs. Octave convolution allows inter-frequency communication between high- and low-frequency bands in latent space. All extracted features are sent to fully connected classifier for element-wise QRS complex detection. Intra- and inter-independent databases testing was conducted to evaluate methods' generalization capacity and noise immunity.

Our experimental evaluations show that our method results in significant F1 gains on multi independent databases. The proposed method results in average 24.23% improvement comparing with P&T algorithm and 2.16% with SENet, which is the first-place-method in China Physiological signal Challenge (CPSC) 2019. Notably, the proposed method shows superior stability to SENet on the whole databases, especially on databases from out-of-distribution sources of the training set.

Keywords—*CNN, ECG, Noise immunity, Octave convolution, QRS complex detection*

I. INTRODUCTION

Electrocardiogram (ECG) records the bio-electrical response of heartbeats and provides an important role in non-invasively monitoring and clinical diagnosis of cardiovascular disease (CVD). A normal heart beat consists of a P wave, QRS complex and a T wave. QRS complex detection is the basis of ECG analysis and promotes the development of automated ECG characteristics analysis, such as R-R interval, heart rate [1], and P-R interval, etc.

Many automated QRS complex detection approaches have been proposed over the last three decades. Due to noise uncertainty and the intra- and inter-patient variability, challenges in QRS detection for wearable ECG still remain. The traditional QRS detection method generally contains the preprocessing stage and the decision stage. The preprocessing stage mainly includes linear or non-linear filter [2] and wavelet transform [3] [4] for denoising in ECG. The decision stage consists of QRS detection and decision logic, including singular value decomposition (SVD) [5] and Hilbert transform [6], etc. Most of these approaches reported

acceptable accuracy on the specific databases for study. However, The QRS detection methods for standard ECGs fail in noisy environments. A recent study [7] confirmed that none of the common QRS algorithms can obtain 80% detection accuracy when tested on a common dynamic noisy ECG database.

In recent years, with increasing computational power, several deep learning methodologies have been utilized for ECG analysis [8]. Recurrent Neural Network (RNN) and Convolution Neural Network (CNN) are the most widely applied. Some neural network (NN)-based ECG analysis methods for different tasks have been proposed, including for disease diagnosis [9], arrhythmia classification [10], sleep staging [11] and biometric human identification [12]. For NNs used in the QRS detection, Xiang *et al.* [13] utilized a two-level CNN followed by a fully connected layer. Yuen *et al.* [14] combined CNN and LSTM to effectively detect QRS complexes in noisy ECGs. In the China Physiological Signal Challenge 2019 (CPSC2019) [15] for QRS detection, Cai *et al.* developed a multi-branch squeeze-and-excitation network (SENet) and achieved the first prize [16].

Despite the NN-based methods that really make a certain level of accuracy in QRS complex detection, some problems still have yet to be solved. The common issue of these supervised approaches is the unsatisfied detection performance on the unknown ECG recordings. Meanwhile, noise in ECG recordings collected from wearable devices has a negative impact on the performance of these methods [17]. For example, the CNN-LSTM method proposed in [14] achieved less than 85% accuracy on the ECGs with a signal-to-noise ratio (SNR) of 0dB.

To address the issues mentioned above, this paper proposes a 20-layer CNN with octave convolution for multi-frequency representation and communication. We evaluate our model on three independent databases: MIT-BIH Noise Stress Test (NST) database [18] [19], QT database [20], and St Petersburg INCART 12-lead arrhythmia database [19]. The input ECGs were pre-processed by discrete wavelet transform. Our model contains two-level CNN branches, which extract features from high- and low- frequency level inputs respectively, then exchange information by octave convolution. Octave convolution is a novel convolution operation proposed by Chen *et al.* [21] for inter-frequency communication and reduction of spatial redundancy while requiring lower computation resource and memories than

* Chengyu Liu is the corresponding author. (e-mail: chengyu@seu.edu.cn).

vanilla convolution. Thus, our proposed method consists of three steps: pre-processing via discrete wavelet transform, element-wise QRS detection via the proposed CNN model, and R peak position determination. The overview diagram is illustrated in Fig. 1.

The rest of this paper is organized as follows. Section II describes the ECG databases used in this work. In Section III, the proposed method is illustrated. Section IV presents the evaluated results. The results are analyzed in Section V. Finally, Section VI concludes the paper. The main contributions of this paper include:

1. We propose a new CNN model for QRS complex detection with inter-patient generalization ability.

2. We investigate the effectiveness of octave convolution on QRS complex detection.

II. DATA

A. Training Data

The training ECG recordings are from CPSC2019 [15]. The publicly CPSC2019 database consists of 2,000 single-lead ECG recordings collected from patients with CVD. Each of the recordings lasts for 10 s. These ECG recordings were obtained from multiple sources and re-sampled to 500 Hz. The R-peak annotations for each recording were given as a position sequence $[r_1, \ldots r_n]$. We split the database recordings from publicly CPSC2019 database into a training and validation set. The R-peaks annotation of each ECG recording was mapped into a binary sequence for supervised training. As the maximum duration of a normal QRS complex is below 0.12 s, 3 timesteps (0.048 s) before and 5 timesteps (0.08 s) after the annotated R peak were set to 1, while the others were set to 0 in the binary sequence. In order to increase the diversity of dataset and avoid overfitting, data augmentation in the training set was applied by adding Gaussian white noises, random pulses, sinusoidal baseline, and sawtooth baseline.

B. Testing Data

Non-publicly CPSC2019 testing database, MIT-BIH NST, QT, and INCART databases were adopted in this work for model testing. The recordings from MIT-BIH NST, QT, and INCART databases were unified to 500 Hz. PhysioToolkit Noise Stress Test [18] adds noise recordings with given SNR to the clean ECG signals. The noise recordings include record *em* (electrode motion artifact), record *ma* (muscle noise) and record *bw* (baseline wander). Record *em* contains significant amounts of muscle noise and baseline wander as well. The noisy wearable ECG devices output ECG signals with SNR in the range of 12 dB to 0 dB [14]. Electrode motion artifact is generally the most troublesome type of noise for ECG analysis due to the similar
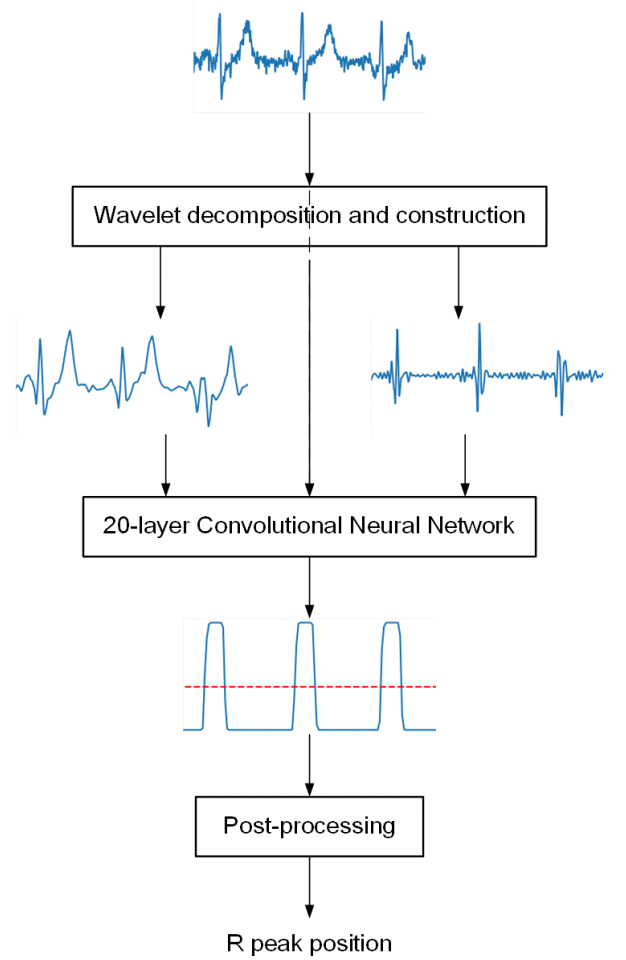
Fig. 1. Overview of our proposed QRS detection method.

characteristic to ECGs. Thus, record *em* with SNR of 0 dB was added to the ECG recordings from MIT-BIH Arrhythmia database [19] [22] to construct the MIT-BIH NST database. The other noise recordings and gaussian white noise were used to measure the robustness to noise of our model and the effectiveness of octave convolution. The databases used in this work are summarized in Table I. We only used the primary ECG lead for QRS complex detection. ECG recordings from each database were split into 10 s ECG signals (i.e. 5000-point ECG sequences) for testing.

III. METHOD

A. Problem Formulation

The QRS detection task takes as input an ECG signal $X = [x_1, \ldots x_t]$, and outputs a sequence of labels $P = [p_1, \ldots p_n]$, such that each p_i represents the probability that a QRS complex appears on the time step. Our model takes the 10 s pre-processed ECG signal sampled at 500 Hz as input

TABLE I. CHARACTERISTICS OF ECG DATABASES USED FOR TESTING IN THIS STUDY

DB Name	CPSC2019-TEST	MIT-BIH NST	QT	INCART
Source Hz	500	360	250	257
Target Hz	500	500	500	500
# Records	1,645	48	105	75
Used # Records	1,645	48	82	75
Records Length	10 s	30 min	15 min	30 min
# Lead	1	2	2	12
Used # Beats	22,609	98,262	78,766	158,366

and outputs a 10 s probability sequence sampled at 62.5 Hz. The output sequence predicts QRS complexes by producing a rectangular function at the location of the R peaks. The reason why the output sequence is down-sampled to 62.5 Hz is that, considering the frequency band of QRS complex (0.5 Hz ~ 25 Hz), a sample rate of 50 Hz is sufficient for QRS complex detection according to Nyquist sampling theorem [23]. ECG input with low sample rate allows our model to focus on the QRS complex and other sub-waves while ignoring high-frequency noise. Meanwhile, as the frequency band of ECG is in range of 0.5 to 100 Hz, the high-frequency information (over 25 Hz) can facilize the recognition of QRS complexes. Our model is designed to fuse the high- and low- frequency information and map features to the low-frequency dimension for QRS detection.

B. Pre-processing and Post-processing

Pre-processing aimed to denoise in ECG signals to prevent our model learning wrong features. Discrete wavelet transform pre-processing was adopted to suppress noise in our method. The 10 s raw ECG signal $X[n]$ was firstly decomposed to generate approximation coefficients (C_1) and detail coefficients (D_1) using Symlets (sym4) wavelet of the first order. C_1 and D_1 were respectively used to reconstruct the approximation component $X_{1,L}[n]$ and the detail component $X_{1,H}[n]$. Then the raw ECG signal and the decomposed components were all fed into the model as inputs. The output of our model was post-processed to obtain the final QRS detection result. A threshold of 0.5 was used for binary classification.

C. Model Architecture

We proposed a 1D CNN with octave convolution, which was shown in Fig. 2. An average pooling operation with a pool size of 8 was performed to generate input for low-frequency level branch, and the high-frequency level branch took as input the pre-processed ECG signal. We arrived at architecture with two 9-layer branches followed by a fully connected layer as the QRS discriminator. Two octave convolution layers were inserted before the 1st and the 10th layers for high- and low-frequency information exchange between branches. The detailed operation of octave convolution was shown in Fig. 3. Inter-frequency communication made the features extracted from each branch
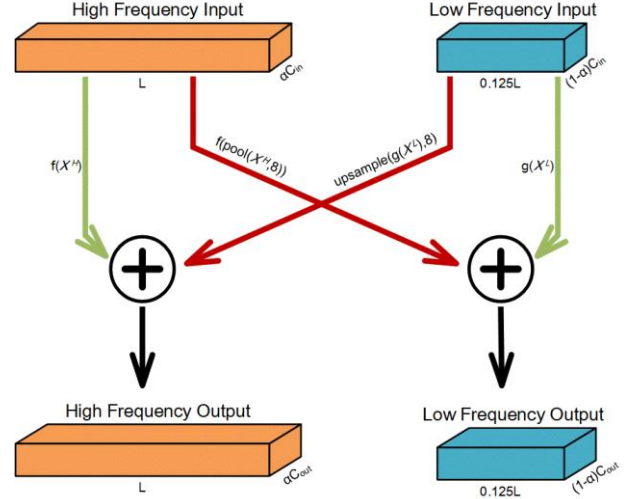
Fig. 3. Detailed operation of octave convolution. C_{in} and C_{out} is the number of filters. In this work, $\alpha = 0.5$.

more robust. Dilated convolution layers were applied in these two branches to enlarge the receptive field through a fewer number of convolution layers. To remove gridding artifacts of dilated convolution, a vanilla convolution layer was applied at the end of each branch. Each branch consisted of 8 dilated convolution layers and a vanilla convolution layer. The dilated convolution layers all had a filter length of 3, 8×2^k filters and a dilated rate of 2^k, where k started as 1 and was incremented every 2-th dilated convolution layer. The octave convolution layers and vanilla convolution layers had the same size of filters as the nearest dilated convolution layer.

After each convolution layer, we applied Batch Normalization [24] and a leaky rectified linear unit (Leaky ReLU). We applied Dropout [25] before the second octave convolution layer where produced the most parameters. The feature maps extracted by the high-frequency level branch were down-sampled by average pooling operation with a pool size of 8, then concatenated with the feature maps extracted by the low-frequency level branch before sending to the fully connected classifier. In the fully-connected classifier, we discarded the flatten operation and directly connect every neuron of each time step with the previous layer's neurons of the same time step by the dense layer. The number of neurons
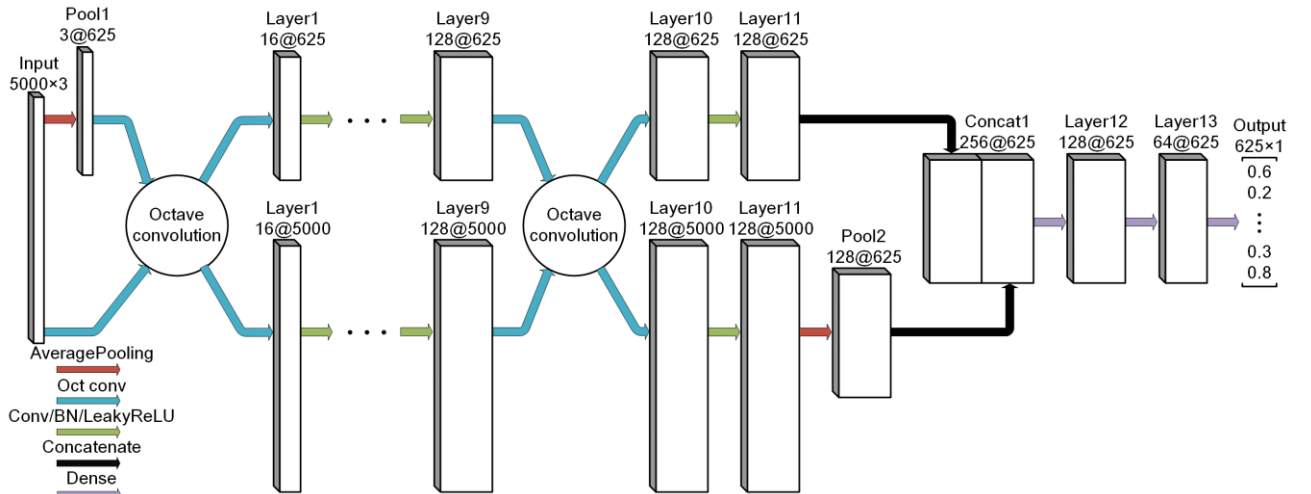
Fig. 2. The architecture of the network. The network contains 2 branches with 9 convolution layers each and 2 octave convolution layers followed by a fully-connected classifier and a sigmoid, then outputs a 625-point sequence with values in the range of [0,1] to predict the probability that a QRS complex appears on each time step.

was decreased layer-wise. Finally, a sigmoid was applied to make a binary prediction for every time step.

IV. EVALUATION METRICS

Evaluated metric contains sensitivity (*Se*), positive precision rate (*PPR*) and *F1*. True positive (*TP*) is the number of correct QRS detection. False positive (*FP*) is the number of incorrect QRS detection. False negative (*FN*) is the number of incorrect rejected QRS. Correct QRS complex detection position must be within 75 ms from the annotated R peak [27]. *TP*, *FP* and *FN* are counted from the whole database. Here, *Se*, *PPR* and *F1* are computed according to the equations below.

$$Se = \frac{TP}{TP+FN} \tag{1}$$

$$PPR = \frac{TP}{TP+FP} \tag{2}$$

$$F1 = \frac{2 \cdot Se \cdot PPR}{Se + PPR} \tag{3}$$

V. RESULT

A. Comparison with Rule-based Algorithms

The testing results of the P&T algorithm [26] and our model on the CPSC2019-TEST, MIT-BIH NST, QT, and INCART database were shown in Table II. The CPSC2019-TEST database was independent of the training database. The P&T algorithm [26] showed the acceptable *F1* score on the CPSC2019-TEST database. However, the *F1* values decreased significantly on the rest databases, where the algorithm reported the lowest *F1* of 0.6103 on MIT-BIH NST. As shown in Table II, the *F1* results of our QRS detection method were higher than 0.95 for most databases except for MIT-BIH NST. The proposed method outperformed the P&T algorithm on all testing databases, noticeably on the QT
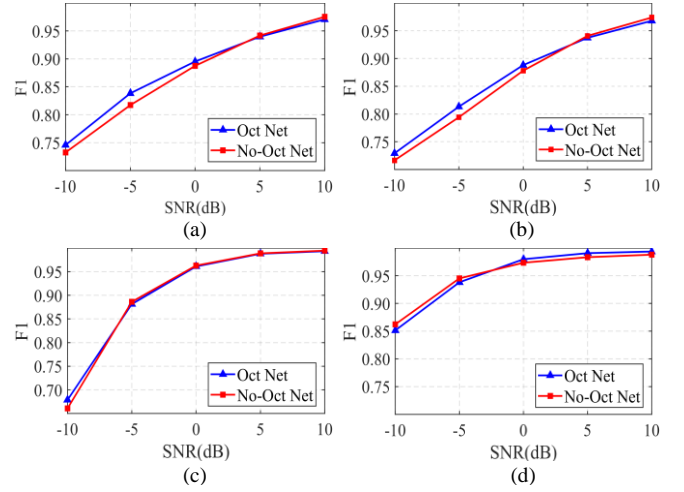
Fig. 4. (a) *F1* score on the MIT-BIH NST database with record *ma*. (b) *F1* score on the MIT-BIH NST database with record *em*. (c) *F1* score on the MIT-BIH NST database with gaussian white noise. (d) *F1* score on the MIT-BIH NST database with record *bw*.

database which includes a wide variety of QRS and ST-T morphologies.

B. Comparison with SENet

SENet for QRS detection was proposed by [17], which won the first prize in CPSC2019. Table II showed that Oct Net significantly reduced the number of *FP* on the whole databases. The performance of Oct Net was more stable than SENet. For the databases expect for CPSC2019-TEST, Oct Net achieved the best performance on *PPR* and *F1* (0.8882 to 0.8541 on MIT-BIH NST, 0.9873 to 0.9581 on QT and 0.9568 to 0.9290 on INCART). Although SENet was more inclined to detect QRS complex, it also caused *PPR* of SENet to deteriorate on most databases, especially on MIT-BIH NST (0.7697 for SENet and 0.8560 for Oct Net). For database CPSC2019-TEST, Oct Net performed slightly worse

TABLE II. PERFORMANCE OF SERVAL QRS DETECTION METHODS ON DIFFERENT DATABASES

Database	Method	*TP*	*FP*	*FN*	*Se*	*PPR*	*F1*
CPSC2019-TEST	P&T algorithm [26]	19,462	4,297	3,147	0.8608	0.8191	0.8395
	SENet [17]	22,396	1,257	213	**0.9906**	0.9469	**0.9682**
	No-Oct Net	21,925	1,008	684	0.9697	**0.9560**	0.9628
	Oct Net (proposed)	22,111	1,179	498	0.9780	0.9494	0.9635
MIT-BIH NST	P&T algorithm [26]	50,116	15,865	48,146	0.5100	0.7596	0.6103
	SENet [17]	94,268	28,202	3,994	**0.9594**	0.7697	0.8541
	No-Oct Net	88,238	14,467	10,024	0.8980	**0.8591**	0.8781
	Oct Net (proposed)	90,681	15,247	7,581	0.9228	0.8560	**0.8882**
QT	P&T algorithm [26]	48,818	14,882	29,948	0.6198	0.7667	0.6853
	SENet [17]	77,767	5,804	999	**0.9873**	0.9306	0.9581
	No-Oct Net	75,853	749	2,913	0.9630	0.9902	0.9764
	Oct Net (proposed)	77,100	318	1,666	0.9788	**0.9959**	**0.9873**
INCART	P&T algorithm [26]	90,103	13,167	68,263	0.5690	0.8725	0.6888
	SENet [17]	153,730	18,879	4,636	**0.9707**	0.8906	0.9290
	No-Oct Net	143,975	5,504	14,391	0.9091	0.9631	0.9352
	Oct Net (proposed)	150,437	5,663	7929	**0.9499**	**0.9637**	**0.9568**

Oct Net is our proposed model. No-Oct Net is the variation of our model which replaced octave convolution layers by vanilla convolution layers.

comparing of the *F1* value with SENet (0.9682 for SENet and 0.9635 for Oct Net).

C. Comparison of Oct Net and No-Oct Net

No-Oct Net was the variation of our model which replaces octave convolution layers by vanilla convolution layers with the same size of filters. As shown in Table II, Oct Net and No-Oct Net both achieved superior performance. The *F1* results as well as the other performance metrics of No-Oct Net were slightly lower than Oct Net on most databases. The performance gap was mainly reflected in *Se* results, around 4% higher for Oct Net than No-Oct Net.

D. Performance on Noisy Databases

Record *ma* and record *em* contain noise with an overlapped frequency band of ECGs. The baseline wanders in the record *bw* is a low-frequency signal. Gaussian white noise has uniform power across the frequency band. Fig. 4. presented the *F1* scores of Oct Net and No-Oct Net on the MIT-BIH NST databases with different types of noise and SNR. The performance of Oct Net and No-Oct Net was reduced with the decreasing of SNR. Oct Net significantly outperformed No-Oct Net on the MIT-BIH NST databases with record *ma* and record *em*, which primarily contain high-frequency noises. Record *bw* and Gaussian white noise had a similar impact on the performance of Oct Net and No-Oct Net.

VI. DISCUSSION

Automated QRS complex detection methods aim to help cardiologists label the vast amounts of ECG data for further ECG analysis. Our method has relatively higher accuracy on various unknown databases and noisy databases, indicating a certain level of generalization ability and robustness. For database CPSC2019-TEST, the proposed method performed slightly worse than SENet, which is the optimal model in CPSC2019. Considering the weight file of SENet had been trained to fit the data distribution of the data source in the competition, this result is acceptable.

Octave convolution reduces computation and memory overhead. Oct Net (0.39 M) only produces 89% of the parameters of No-Oct Net (0.44 M) and reduces the FLOPs of No-Oct Net by 30%. Moreover, Oct Net shows superior performance to No-Oct Net on ECG recordings contaminated by high-frequency noise. The testing results on the MIT-BIH NST database with different types of noise recordings demonstrate octave convolution tends to suppress the high-frequency noise and maintain the performance when dealing with low-frequency noise.

VII. CONCLUSION

This paper has presented an octave convolution-based QRS complex detection method. Two CNN branches are designed to extract features from multi-frequency ECG signals. Octave convolution allows inter-frequency communication of multi-frequency ECG signals' feature maps while requiring less computation and memory overhead than vanilla convolution. The testing results obtained by using intra- and inter- database testing demonstrate that this proposed method provides a higher generalization ability and robustness to noises than other NN-based and traditional state-of-art QRS detection methods. As a result, the proposed method has the potential for application in the noisy daily environment. In future work, noise in the frequency band of QRS complex needs to be removed for more accurate feature extraction. And other biomedical signals will be combined with the ECG signals to improve the accuracy and robustness of QRS detection.

REFERENCES

[1] Y. T. Xie, J. Q. Li, T. T. Zhu and C. Y. Liu, "Continuous-Valued Annotations Aggregation for Heart Rate Detection," *IEEE Access*, vol. 7, no. 1, pp. 37664-37671, 2019.

[2] Z.E.H. Slimane and N.A. Amine, "QRS complex detection using empirical mode decomposition," *Digit Signal Process*, vol. 20, no. 4, pp. 1221–1228, 2010.

[3] M. Alfaouri and K. Daqrouq, "ECG signal denoising by wavelet transform thresholding," *American Journal of Applied Sciences*, vol. 5, no. 3, pp. 276–281, 2008.

[4] S. Kadambe, R. Murray, and G. F. Boudreaux-Bartels, "Wavelet transformed-based QRS complex detector," *IEEE Transactions on Biomedical Engineering*, vol. 46, pp. 838–848, 1999.

[5] W. H. Jung and S. G. Lee, "An R-peak detection method that uses an SVD filter and a search back system," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 3, pp. 1121–1132, 2012.

[6] M.S. Manikandan, K.P. Soman, "A novel method for detecting R-peaks in electrocardiogram (ECG) signal," *Biomedical Signal Processing and Control*, vol. 7, no. 2, pp. 118–128, 2012.

[7] F. Liu, C. Liu, X. Jiang, Z. Zhang, Y. Zhang, J. Li, and S. Wei, "Performance analysis of ten common QRS detectors on different ECG application cases," *Journal of Healthcare Engineering*, vol. 2018, no. 2, pp. 9050812(1)–9050812(8), 2018.

[8] C. Y. Liu, M. C. Yang, J. N. Di, Y. T. Xing, Y. W. Li and J. Q. Li, "Wearable ECG: History, Key Technologies and Future Challenges," *Chinese Journal of Biomedical Engineering*, vol. 38, no. 6, pp. 641-652, 2019.

[9] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals," *Information Sciences*, vol. 416, pp. 190–198, 2017.

[10] Y. H. Awni, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, pp. 65–69, 2019.

[11] J. Malik, Y. L. Lo and H. T. Wu, "Sleep-wake classification via quantifying heart rate variability by convolutional neural network," *Physiological Measurement*, vol. 39, no. 8, 2018.

[12] M. Hammad, Y. Liu, and K. Wang, "Multimodal biometric authentication systems using convolution neural network based on different level fusion of ECG and fingerprint," *IEEE Access*, vol. 7, pp. 26527–26542, 2019.

[13] Y. Xiang, Z. Lin and J. Meng, "Automatic QRS complex detection using two-level convolutional neural network," *BioMedical Engineering OnLine*, vol. 17, no. 1, pp. 13, 2018.

[14] B. Yuen, X. Dong, T. Lu, "Inter-Patient CNN-LSTM for QRS Complex Detection in Noisy ECG Signals," *IEEE Access*, vol. 7, pp. 169359–169370, 2019.

[15] H. Gao, C. Liu, X. Wang, L. Zhao, Q. Shen, E. Y. K. Ng, and J. Li, "An open-access ECG database for algorithm evaluation of QRS detection and heart rate estimation," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 9, pp. 1853–1858, 2019.

[16] W. Cai, D. Hu, "QRS complex detection using novel deep learning neural networks," *IEEE Access*, vol. 8, pp. 97082–97089, 2020.

[17] C. Y. Liu, X. Y. Zhang, L. N. Zhao, F. F. Liu, X. W. Chen, Y. J. Yao and J. Q. Li., "Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system," *IEEE Internet of Things Journal*, vol. 2, no. 2, pp. 1363-1374, 2019.

[18] G. B. Moody, W. E. Muldrow, and R. G. Mark, "A noise stress test for arrhythmia detectors," *Computers in Cardiology*, Los Alamitos, CA: IEEE Computer Society Press, vol. 11, pp. 381–384, 1984.

[19] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215-e220, 2000.

[20] P. Laguna, R. G. Mark, A. L. Goldberger, and G. B. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," *Computers in Cardiology*, Los Alamitos, CA: IEEE Computer Society Press, pp. 673–676, 1997.

[21] Y. Chen, H. Fan, B. Xu, Z. Yan, Y. Kalantidis, M. Rohrbach, S. Yan,

and J. Feng, "Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3435–3444, 2019.

[22] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database", *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[23] E. Ajdaraga, M. Gusev, "Analysis of sampling frequency and resolution in ECG signals," *Telecommunication Forum*, pp. 1–4, 2017.

[24] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[25] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[26] J. Pan, and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, pp. 230–236, 1985.

[27] Association for the Advancement of Medical Instrumentation, "Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms," *American National Standards Institute, Inc.*, ANSI-AAMI:EC57, 1998.

Capacitance Measurement of Molten Metal Level in Continuous Casting System

Zhe Zhuang
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an China
ORCID: 0000-0003-1275-5112

Yu Zhang
Management Department
Shanghai Hin*ner Information*
Technology Co., LTD
Shanghai, China
447820892@qq.com

Tonghao Zhou
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an China
ORCID: 0000-0002-6904-7056

Shengchang Ji*
State Key Laboratory of
ElectricalInsulation and Power
Equipment,School of Electrical
Engineering
Xi'an Jiaotong University
Xi'an,China
ORCID: 0000-0001-9293-9549

Abstract—**The accuracy of molten metal level measurement is the base of controlling the molten metal level, which greatly affects the production quality. Due to the oscillation of the caster mold, hot temperature, and noises in the continuous casting system, the accurate level is hard to obtain. Currently used molten metal level measurement technique includes radioisotope transmission, electromagnetic, ultrasound reflection, microwave reflection, computer vision, and capacitive level measurement method. The radioisotope transmission method has potentially hazardous to exposed workers, while the electromagnetic method needs to maintain the coil frequently for the large excitation source needed, the ultrasound reflection method, microwave reflection method and computer vision method can be affected by the casting powder used for preventing molten metal splash. In this work, a new structure of the capacitance sensor designed for minimizing the influence of caster mold's oscillation and improving heat dissipation was proposed, the variation of capacitance between the sensor electrode and the caster mold with molten metal level change and the oscillation of the caster mold was calculated in simulation and validate through a grounded simulation tank with capacitance to digital converter FDC1004. The result shows that the proposed capacitance sensor can achieve an accuracy of 1 mm when the distance between the molten metal surface to the sensor within 80 mm with 2 samples per second report speed.**

Keywords—continuous casting, molten metal, level sensor, capacitance sensor, directional enhancement

I. INTRODUCTION

Continuous casting is a technology to realize continuous metal slab casting, which greatly improved the productivity and quality of metal slabs compared to older metal molding technology. The flow velocity, impurities, and unstable metal oxides greatly involved the quality of the produced metal slab[1]–[4]. In order to improve the quality and product stability of the produced metal slab, one approach is to control the flow rate during the process, controlling the level of the molten metal level at the mold[1], [5]. As an important part of controlling the molten metal level, the accuracy and sample speed of the molten metal level sensor had a great influence on the control precision and response speed of the molten metal level control system.

Fig. 1 shows the typical structure of a working caster mold, two copper-walled water tanks are used to cool down the molt-
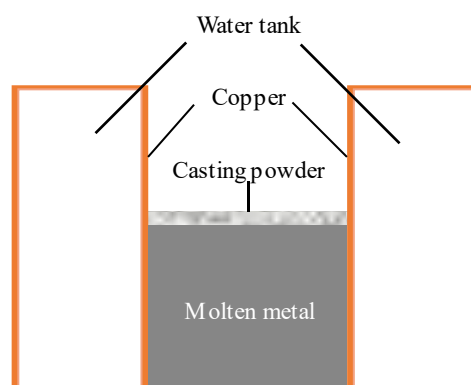
Fig. 1. The typical structure of a working caster mold

en metal between water tanks, while there is a layer of manually added casting powder to improve the product quality. The target is to control the level of molten metal in the caster mold. However, for the casting powder is added manually, the thickness of casting powder is not well-distributed, which can influence the measure of the molten metal level. During the process of continuous casting, the two water tanks are oscillating to prevent molten metal from adhering to the copper wall of the water tank[1]. Due to the high conductivity and high temperature of molten metal in continuous casting system, some traditional contact liquid level measurement methods[6]–[8] are difficult to apply. At present, the methods for measuring the level of molten metal include radioisotope transmission[9], [10], ultrasonic pulse[11], [12], laser[12]–[14], computer vision process[15], [16],thermal[15], [17], capacitance measurement[18] and electromagnetic measurement[19]–[21]. Due to the radiation leakage caused by radiation penetration will affect the human body, it is rarely used in production environments with human participation. The ultrasonic pulse and laser methods can only measure a tiny area, which means a matrix of sensors are needed to get the overall level of the caster mold, while the measured level includes the layer of casting powder, the molten metal level is hard to obtain. The computer vision method can also be affected by the casting powder layer, but with the assistant of thermal level sensing technology, the molten metal level can be obtained[15]. However, the accuracy of the computer vision level measurement is lower than other methods, it's about 5 mm. The electromagnetic method using two coils to test the reluctance of air, caster mold's water tank, and molten metal, which can achieve good linearity in the measuring

---

range. But for the caster mold's water tank is included, the measurement result of electromagnetic method can be affected by the caster mold's oscillation. The capacitive method is sensitive to the surrounding objects, as the distance of the sensor and molten metal surface is long, the influence of the environment is large.

In this work, the directionality of the capacitive sensor was enhanced to minimize the influence of surrounding objects and caster mold's oscillation. A simulation of the capacitance variation was done by COMSOL Multiphysics, the results were verified through a copper wall steel surface water tank.

II. CAPACITANCE SENSOR DESIGN

The structure of the designed sensor is shown in Fig. 2. The simulation model of the sensor is implemented with a print circuit board, which consists of a circular sensor electrode, the remote ground, and a barrel of active-shielding electrode. The two electrodes are both made of copper to ensure the conductivity of the electrode, while there are many metalized holes applied to the shielding electrode to ensure the equipotential of the shielding electrode and improve the heat dissipation capability with running cold air provided at the metal factory. The radiation of the sensor electrode is 15 mm. In order to enhance the shielding effect and the directionality of the sensor, the inner radius of the shield electrode is 16 mm and the outer radius is 50 mm.
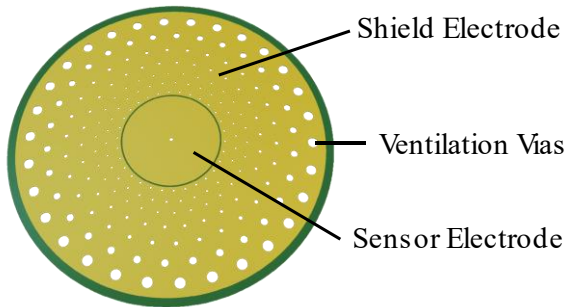
Fig. 2. The structure of designed capacitive sensor

Compared with the previous designed molten metal level sensors used in the caster mold, the proposed sensor design can reduce the influence of caster mold's oscillation. For electromagnetic sensor, because the copper wall of the caster mold has good conductivity, and the upper part of the caster mold wall is fixed by the iron frame which has high permeability, the mold wall played an important role in the magnetic circuit, resulting in the measurement results containing a large mold wall vibration signal. For normal capacitance sensor which only shielded the connection line of converter board and sensor, due to the open structure of the mold, not only the position of the conductive mold wall will affect the capacitance value between the sensor and the ground, but also the surrounding human and other conductive bodies will affect the measurement results. In this design, the shielded sensor electrode can be focused on the part directly opposite the molten metal surface, thereby reducing the influence of irrelevant surrounding objects.

The capacitance is obtained through Texas Instruments' capacitance to digital converter FDC1004, which includes a shield output port to reduce the influence of stray capacitance, however, the buffer included in the chip can only shield 400pF capacitance at most, a standalone shield buffer is needed to shield large stray capacitance in the continuous casting

operation. The capacitance measurement circuit is shown in 0, an additional buffer is applied to the chip's shield port to enhance the shielding capacity of the converter. The capacitance converter and the caster mold were well grounded, the converter measures the capacitance between the sensor electrode and the ground. The capacitance converter uses a 2.9 V DC voltage source as the excitation, while the embedded capacitor and the external capacitor formed a charge distribution circuit. When the connection between the embedded capacitor and the external capacitor is disconnected and the embedded capacitor is discharged, the external capacitor is charged by the excitation source given by the converter chip, then the excitation source is removed and the external capacitor is connected to the embedded capacitor. After the charge distribution process is complete, the value of the extern capacitor can be captured through the measurement of the embedded capacitor voltage. In this case, a buffer following the output of the sensor can shield the stray capacitance irrelevant to the measurement.
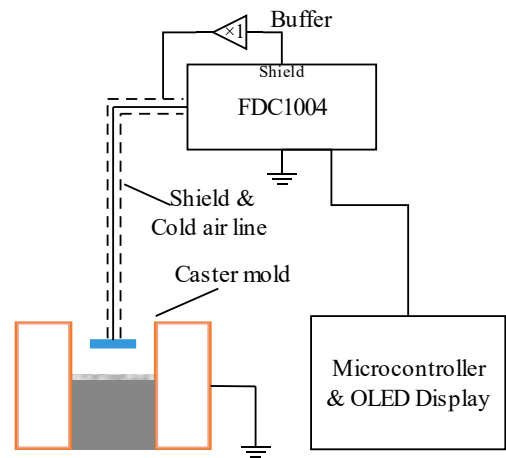
Fig. 3. Designed capacitance measurement circuit

III. SENSOR SIMULATION

According to the targeted caster mold dimensions, a module of the caster mold was designed. The total length of the caster mold is 1.1 m, and the width of the caster mold is 0.2 m. The measurement range required to sense the molten metal level should be greater than 0.06 m. Considering the measurement principle of the capacitance converter, the capacitance sensor electrode is simplified to a shielded circular plate. The central part of the simulated sensor is shown in Fig. 4 (a). The shielding layer covered the sensor electrode and applied to the same direct current voltage source. Therefore, if only by monitoring the charge of sensor electrode, the calculated capacitance can be equivalent to the capacitance in actual operation. For the length of the caster mold is relatively long, the influence of the two copper wall terminals are not considered in the simulation module, only molten metal surface and two copper walls of the caster mold water tank are considered. The complete simulation module built in COMSOL Multiphysics is shown in Fig. 4(b), the three cuboid surfaces facing the sensor are applied to the ground. For the variation of capacitance between

the sensor electrode and the ground is quite small, the relative tolerance of the solver is set to 10^{-7}.

(a)

(b)
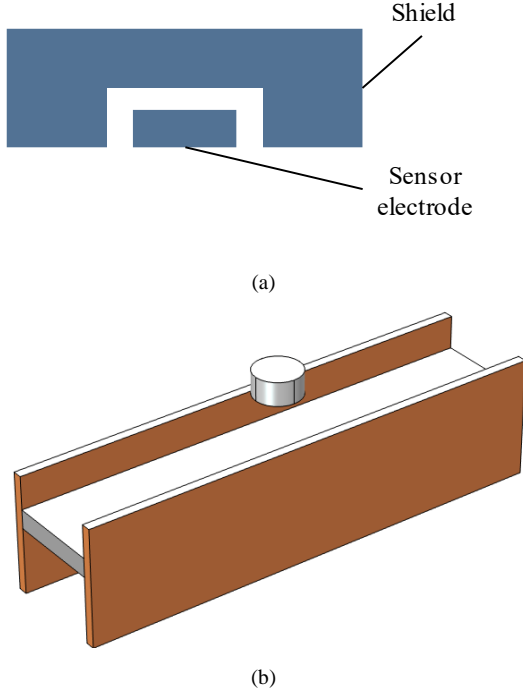
Fig. 4. Designed COMSOL simulation module:(a) Center section of the capacitance sensor.(b) Complete simulation module

Considering the initial distance of the capacitive sensor to molten metal surface needed to ensure the temperature of the sensor is not too high, the initial distance of the capacitive sensor and the molten metal surface is 0.051m, at the initial state, the surface of the molten metal was as high as the copper wall of the caster mold's water tank. The simulation result of the capacitance measured between the sensor electrode and ground varies with different caster mold position and the molten metal level is shown in Fig. 5, the distance is of the sensor electrode to molten metal surface. Corresponding to the simulation result, without caster mold's position correction, the error of capacitance measurement is smaller than 1mm while the distance of the sensor electrode and molten metal surface within 0.1 m which shows the influence of the caster mold's oscillation to the capacitance measurement result is small, the shield designed in the sensor works well in the simulation. The influence of caster mold's oscillation is increasing when the distance of the sensor electrode and molten metal surface increased shows the shielding effect of the sensor's shield electrode is decreasing when distance increased. Using MATLAB curve fitting toolbox, the most suitable expression format was the power with constant addition, the fitting expression for the caster mold at original position was:

$$C_S = 0.073 \cdot d^{-2.25} + 262.4 \text{ fF} \qquad (1)$$

Where C_S is the calculated capacitance through simulation, d is the distance between the capacitive sensor and the molten metal surface. The R-square of the fitting result was 0.9998, which means the fitting expression highly fits the raw data.

Compared with the capacitance expression of the ideal parallel plate capacitor, the simulation result turns out that the change in capacitance value with distance is power of -2.25, while in an ideal parallel plate capacitor is the power of -1. The reason for this difference is not only the area of the copper wall facing the capacitive sensor varied with molten metal

level, but the large distance between capacitive sensor to the grounded surface caused obvious edge effect. Considering the measurement range of the capacitance converter FDC1004, ±15 pF and the measurement accuracy of the converter was 0.5 fF, the capacitance variation was well within the capacitance converter's measurement range, and the change of capacitance can be clearly identified.
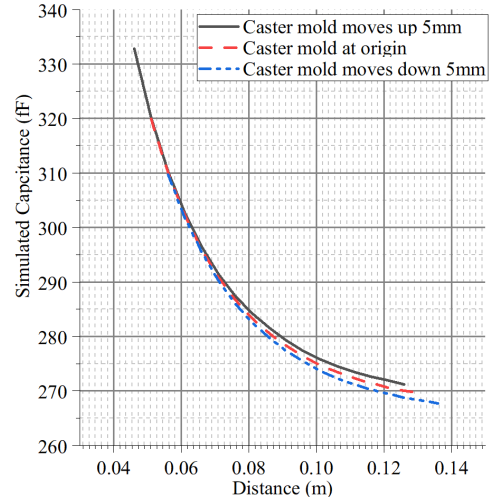
Fig. 5. Capacitance varied with caster mold position and molten metal level

IV. EXPERIMENTAL SETUP
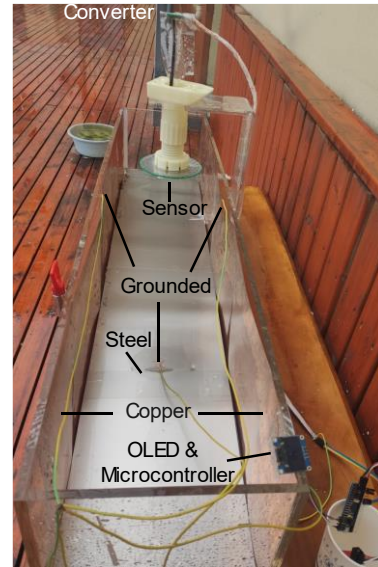
Fig. 6. Simulated molten metal level measurement test system

A simulated molten metal level measurement test system was implemented with a copper wall steel surface acrylic water tank, as shown in Fig. 6, the inner length of the tank is 1.1 m, along with 0.2 m width and 0.25 m height. The capacitive sensor is fixed by an adjustable bracket, Organic Light-Emitting Diode(OLED) display panel and microcontroller was powered by an ungrounded power source and connected to the capacitance converter through shielded cable, while the sensor connects to the capacitance converter through a shielded coaxial cable. The copper walls and the steel surface were grounded through wire connection to the power source's ground to reduce the influence of the remote ground. The oscillation of the caster mold can be simulated by varying the height of the sensor using the adjustable bracket equivalently. The steel surface is made of a layer of foam underneath

bonded with a thin steel plate; the two copper walls were made of 0.2 mm copper foil and pasted at the wall of the acrylic water tank. The level of the steel surface can be varied by adjusting the water level in the acrylic water tank, the foam holds the thin steel plate to float on the water and release the water with a tap installed on the side of the water tank. As the principle of capacitance measurement is to measure the capacitance between the sensor electrode and the ground electrode, which can be calculated through facing area, using copper foils and a thin steel plate to simulate the copper wall water tanks and the molten metal level was reasonable.

If the power source is turned on, the microcontroller will communicate with the capacitance converter and OLED display using Inter-Integrated Circuit Bus(I2C) interface at 50kHz to ensure data transmission reliability, configuring the capacitance converter to 100 samples per second(SPS), single-channel mode, averaging 50 samples to get more stable capacitance and plot the average result to OLED display. As a result, the measured capacitance value can be read easily through OLED display, the sample rate of the molten metal level measurement system is 2 SPS.

V. Results and Discussion

As an initial measurement setting, the distance between the capacitance sensor to the water tank's upper surface is set to 51 mm by adjusting the bracket fixing the capacitive sensor, which diameter was the same in simulation settings. The measurement result of capacitance varying with the distance between the capacitance sensor and the thin steel plate was shown in 0, the trends of capacitance were similar to the simulation result at the same condition, the influence of caster mold's oscillation was relatively small in the distance of 0.04 m to 0.07 m, and gradually increased at longer distance for the caster mold moves down 5 mm curve, while the caster mold moves up 5 mm curve have little difference of the caster mold at origin one, this can be caused by the not fully fitted copper foil at caster mold's side wall, which changed the initial relative place of sensor and caster mold wall. However, the measured capacitance data were much larger than simulation. Fitting the curve through curve fitting toolbox on MATLAB, the fitting expression was:

$$C_S = 0.087 \cdot d^{-2.13} + 1538 \text{ fF} \tag{1}$$
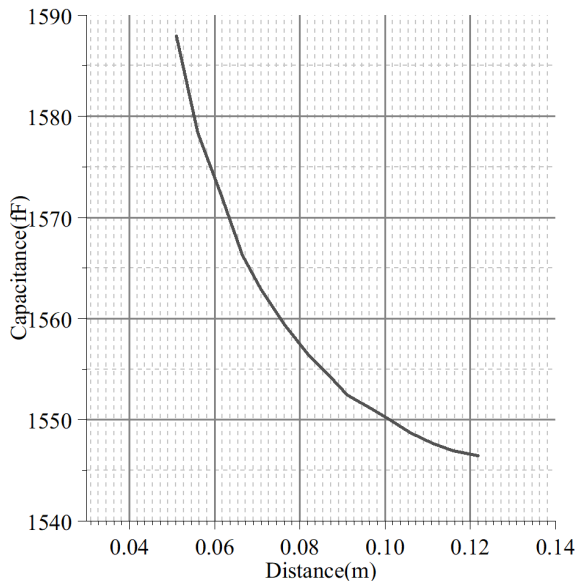
Fig. 7. Measured capacitance value varied with distance at initial settings

Where C_M is the measured capacitance, d is the distance between capacitance sensor and thin steel plate, the R-square of the fit is 0.9996.

There is minor difference in the exponential value and co-efficient of the exponential terms compared to the simulation result, but the value of the additional constant term is of significant difference, which can also be directly seen from the raw data of measured capacitance, the value of capacitance at each point is much larger than those in the simulation result. The larger additional constant can be the stray capacitance in the experimental circuit. By removing the copper foil and thin steel plate's ground connection, the measured capacitance value is 1507 fF, which can prove that the change of additional constant term is due to the stary capacitance of circuit board layout, the wiring of data transmitting lines, and internal circuit of the converter.
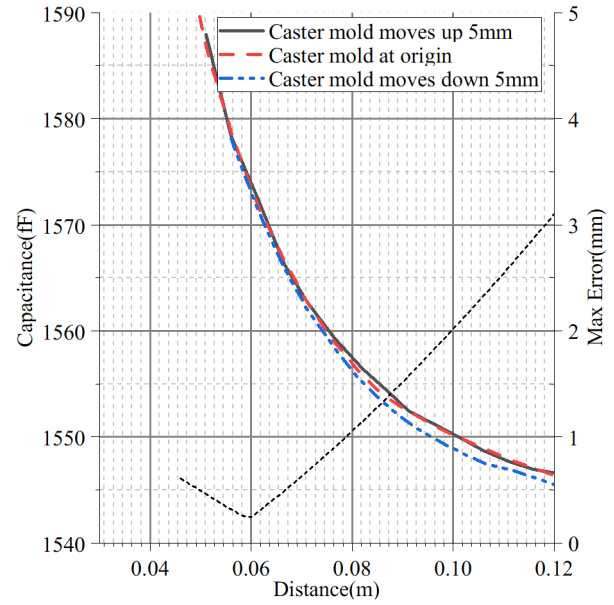
Fig. 8. Capacitance and error varied with distance

By changing the distance between the capacitive sensor and upper surface of the water tank, the caster mold's oscillation was simulated. The error of level measured caused by caster mold's movement varies with the distance between the capacitive sensor and the thin steel plate was derived. The influence of caster mold's oscillation is shown in Fig.8 , the accuracy of the designed sensor can achieve 1 mm within the distance of 80 mm and 2 mm within the distance of 100 mm.

VI. Conclusion

The molten metal level sensor based on the capacitance measurement technique proposed in this work is able to shield the influence of caster mold's oscillation and enhance the directional of capacitance sensing, having the potential to replace current technologies used in molten metal level sensing for higher accuracy and faster sample rate. The experimental result of the designed sensor showed the measurement range of the sensor is 80 mm at the accuracy of 1 mm with the peak-to-peak value of caster mold's oscillation, the change of capacitance matched well with the simulation result obtained through COMSOL Multiphysics. The proposed sensor is still under testing for long term accuracy, and stability on the simulation water tank. Some optimizations of the sensor to suit the temperature, the installation position, measurement accu-

racy and the electromagnetic environment of continuous casting are still in progress. Further researches of the capacitive sensor design to strengthen directivity, propose specific diameter design rules, increase the sensitivity of the level measurement are planned to enhance the performance of designed capacitive sensors. Due to the improved directivity of the capacitive sensor, the design route can also be applied to gesture recognition, object motion detection, and more applications.

REFERENCES

[1] M. Kawamoto and T. Nagira, "Flowering of continuous casting process for steel in japan and new fundamental seeds to the future," *Tetsu--HaganeJournal Iron Steel Inst. Jpn.*, vol. 100, no. 4, pp. 472–484, 2014, doi: 10/gg4m52.

[2] A. Cwudziński, "Physical and mathematical simulation of liquid steel mixing zone in one strand continuous casting tundish," *Int. J. Cast Met. Res.*, vol. 30, no. 1, pp. 50–60, Jan. 2017, doi: 10/gg4m5z.

[3] B. Y. Jiang, L. G. Meng, B. Ya, B. W. Zhou, and X. G. Zhang, "Study on the horizontal continuous casting of Cu-based bulk metallic glass slab," *J. Non-Cryst. Solids*, vol. 543, p. 120150, Sep. 2020, doi: 10/gg4m53.

[4] R. I. L. Guthrie, "A Review of Fluid Flows in Liquid Metal Processing and Casting Operations," *ISIJ Int.*, vol. 49, no. 10, pp. 1453–1467, 2009, doi: 10/b6cgr4.

[5] J. Saleil, M. Mantel, and J. Le Coze, "Stainless steels making: History of production processes developments. Part III. Casting methods, hot and cold forming processes," *Mater. Tech.*, vol. 108, no. 1, p. 105, Jun. 2020, doi: 10/gg4m5x.

[6] Z. Niu, Y. Zhao, B. Tian, and F. Guo, "The novel measurement method of liquid level and density in airtight container," *Rev. Sci. Instrum.*, vol. 83, no. 12, p. 125108, Dec. 2012, doi: 10/gbcddq.

[7] F. Reverter, X. Li, and G. C. M. Meijer, "Liquid-level measurement system based on a remote grounded capacitive sensor," *Sens. Actuators Phys.*, vol. 138, no. 1, pp. 1–8, Jul. 2007, doi: 10/frdhvq.

[8] B. Yun, N. Chen, and Y. Cui, "Highly Sensitive Liquid-Level Sensor Based on Etched Fiber Bragg Grating," *IEEE Photonics Technol. Lett.*, vol. 19, no. 21, pp. 1747–1749, Nov. 2007, doi: 10/fghzxn.

[9] G. Feiyan and H. Ying, "A novel isotopic mould measurement system for molten steel level based on real time algorithm," *J. Test Meas. Technol.*, vol. 22, no. 6, pp. 505–9, Jun. 2007.

[10] M. Köhler, M. Hult, D. Arnold, M. Laubenstein, and J.-L. Reyss, "Reference measurements of low levels of 60Co in steel," *Appl. Radiat. Isot.*, vol. 61, no. 2, pp. 207–211, Aug. 2004, doi: 10/btn6mf.

[11] O. G. H. Nygaard and K. S. Mylvaganam, "Ultrasonic time-domain reflectometry for level measurement in molten metals/ Ultraschall-Impulsreflektrometrie zur Niveaumessung von Metallschmelzen," *Tm - Tech. Mess.*, vol. 60, no. 1, pp. 4–14, Jan. 1993, doi: 10/gg4tjw.

[12] K. Masakazu and M. Toshihiko, "Method For Controlling Casting In Continuous Casting Machine." Japan, Jan. 23, 1996.

[13] Z. Su, Q. He, and Z. Xie, "Indirect measurement of molten steel level in tundish based on laser triangulation," *Rev. Sci. Instrum.*, vol. 87, no. 3, p. 035117, 2016, doi: 10/f8hn87.

[14] L. M. Zhao, Q. Ouyang, D. F. Chen, and L. Y. Wen, "Surface defects inspection method in hot slab continuous casting process," *Ironmak. Steelmak.*, vol. 38, no. 6, pp. 464–470, 2011, doi: 10/b9j9jw.

[15] J. Liu, Z. W. Hu, J. B. Lei, and Z. Xie, "New Solution to Molten Steel Level Measurement in Continuous Casting Tundish," *Acta Electron. Sin.*, vol. 38, no. 5, pp. 1196–1200, 2010, doi: 10/bxgbj3.

[16] Z. Hu, Z. Xie, Y. Ci, and W. Wei, "Molten Steel Level Measuring Method by Thermal Image Analysis in Tundish," in *Recent Advances in Computer Science and Information Engineering: Volume 6*, Z. Qian, L. Cao, W. Su, T. Wang, and H. Yang, Eds. Berlin, Heidelberg: Springer, 2012, pp. 361–367.

[17] T. D. Kaiser, S. S. Daniel, and C. D. Dykes, "Method for determining molten metal pool level in twin-belt continuous casting machines," 1989.

[18] Hü. Canbolat, "A Novel Level Measurement Technique Using Three Capacitive Sensors for Liquids," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 10, pp. 3762–3768, Oct. 2009, doi: 10/dpxsp5.

[19] A. Saleem, P. R. Underhill, D. Chataway, T. Gerritsen, A. Sadri, and T. W. Krause, "Electromagnetic Measurement of Molten Metal Level in Pyrometallurgical Furnaces," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 3118–3125, Jun. 2020, doi: 10/gg35wn.

[20] R. A. Katankin and A. D. Pokrovskiy, "A study of the eddy-current method of testing the level of molten steel in the thin-slab crystallization process," *Russ. J. Nondestruct. Test.*, vol. 45, no. 8, p. 542, Dec. 2009, doi: 10/dp9g35.

[21] R. I. Khalilov, S. Yu. Khripchenko, P. G. Frik, and R. A. Stepanov, "Electromagnetic measurements of the level of a liquid metal in closed volumes," *Meas. Tech.*, vol. 50, no. 8, pp. 861–866, Aug. 2007, doi: 10/ftwjcp.

Application of Wire Mesh Sensor to Identify Gas-liquid Flow Patterns in a Horizontal Pipe

Shuai Liu
School of Nuclear Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
liu-shuai@sjtu.edu.cn

Li Liu[*]
School of Nuclear Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
liulide@sjtu.edu.cn

Jiarong Zhang
School of Nuclear Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
28315286@qq.com

Hanyang Gu
School of Nuclear Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
guhanyang@sjtu.edu.cn

Qi Zhang
School of Nuclear Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
m18846172005@163.com

Abstract— The flow pattern recognition and phase interface imaging are the basis of studying gas-liquid two-phase flow. It can directly reflect the dynamic development process of phase distribution in the flow process, and it is also the basis for theoretical analysis and physical modeling. In order to identify the characteristics of gas-liquid flow patterns by self-developed wire mesh sensor (WMS), the phase concentration imaging analysis of both non-swirling gas-liquid flows and swirling gas-liquid flows induced by a swirler in a 30 mm inner diameter pipeline is carried out and compared with high-speed photography. Tap water and air are used as working fluids. The results indicate that the WMS can reflect the phase concentration distribution of bubble, plug, slug flow, swirling gas column flow and intermittent flow. Especially, the WMS has good imaging ability for large-scale phase interface distribution. To some extent, it can reflect the void fraction characteristics in a pipe, but it cannot reflect the swirling characteristics of swirling gas-liquid flow. The WMS does not have the ability to directly identify the thin liquid film and microbubbles, which needs to be further optimized.

Keywords—self-developed wire-mesh sensor, high-speed photography, flow pattern recognition, gas-liquid two-phase flow, direct imaging.

I. INTRODUCTION

Gas-liquid two-phase flow widely exists in the field of the process industry[1-4]. As the basic research of gas-liquid flow, flow patterns and phase concentration distribution are the important basis for establishing a theoretical model and regulating process operation parameters[5]. The commonly used methods of flow patterns recognition and cross-section phase concentration distribution measurement include high-speed photography visualization, ray irradiation, conductivity probe, tomography, etc. The characteristics of the above methods have their limitations[6-9].

The wire mesh sensor (WMS) is an invasive flow parameter measurement and imaging sensor which can simultaneously measure the flow cross-section of two-phase flow[10,11]. It mainly includes two measuring principles: conductance type and capacitance type, which are mainly selected according to the electrical characteristics of the measuring fluid. Compared with other measurement methods, the measurement accuracy of WMS is relatively high. Its unique advantage is that it can realize the real-time imaging of

the flow section while measuring the flow parameters, which makes the two-phase flow characteristics more intuitive[12]. Therefore, WMS is an important measurement method for the study of gas-liquid flow, which is of important to enrich the effective measurement of two-phase flow parameters.

The WMS was first used to measure the water content of crude oil[13], but it did not involve the application of phase interface imaging. At present, the most researched and widely used WMS is a conductance wire mesh sensor developed by the German scholar Prasser et al[14]. It is applied to the measurement of gas-liquid flow. Its principle is to determine the flow characteristics of the measured fluid by measuring the relationship between the signal and the conductivity of the medium and to observe the fluid distribution of the cross-section by imaging[14]. After that, Prasser and others have done more in-depth application research and optimization for the wire mesh sensor, and have made a wealth of research results[15-20].

Relevant researchers applied WMS to the research of bubble flow and slug flow[21,22]. They found that WMS imaging technology can be applied to large-scale Taylor bubbles and bubbles accumulation area. The invasion effect of WMS and the change of flow patterns are also studied, and the comparison between WMS and traditional conductivity probe sensors in the measurement of the flow parameters is carried out, and it is found that the measurement results were basically consistent[23]. Besides, WMS imaging technology is also used to study the three-dimensional shape of bubbles, gas-liquid interface, gas-liquid concentration distribution, and other fields. At the same time, the adaptability of the WMS in the different shapes of the interface channel is also studied. Based on conductivity WMS, capacitance WMS has also been widely used in the field of multiphase flow, which weakens the electrical limit of multiphase flow, that is, conductivity WMS is selected for the high conductivity fluid medium. For non-conductive fluid medium, the capacitive WMS is selected, which uses the dielectric constant characteristics of the fluid medium to collect electrical signals and conduct data processing and imaging[23-25].

Although the WMS has been widely used in the field of gas-liquid flow, there are few kinds of researches on the adaptability of gas-liquid flow patterns of WMS, especially in the swirling gas-liquid two-phase flow. Based on the visualization flow patterns of high-speed cameras, this paper analyzes the imaging process of WMS in various typical non

---

* Li Liu is the corresponding author. (e-mail: liulide@sjtu.edu.cn).

rotating and swirling gas-liquid two-phase flow patterns, expecting to provide necessary data support for WMS structure design and optimization of data processing methods.

II. EXPERIMENTS AND MEASURING SYSTEM

A. Experiment flow chart

As shown in Fig. 1, the experimental system includes four main parts: the liquid circuit, gas circuit, high-speed photography system, and the WMS system. The pipe material is plexiglass with an internal diameter of 30 mm. The experimental medium is tap water and compressed air under normal temperature and pressure. The gas and liquid superficial velocities are j_g = 0.016~21.357 m/s and j_l = 0.020~2.269 m/s, respectively. A swirler is installed 2000 mm downstream of the outlet of the gas-liquid mixer, which consists of four helical blades with a 180° twist, and the outlet angle is 30°. Its function is to change the non-swirling into swirling flow. At the same time, the WMS and high-speed camera record the non-swirling and swirling flow state at the upstream and downstream z/D = 10 of the swirler. In the process of measurement, the position of the high-speed camera is set at the upstream of the WMS. Because they are very close to each other, the flow patterns almost remain the same, so it can be approximately considered that both of them can imagine the same position of the pipe.
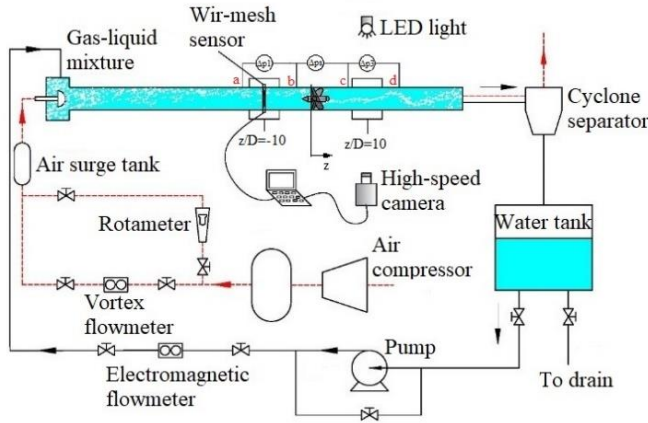
Fig. 1. Diagram of the experiment setup.

B. Wire Mesh Sensor
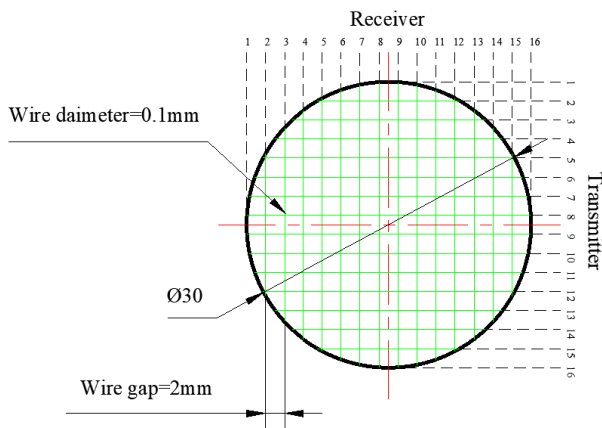
Fig. 2. The structure design of WMS.

The arrangement of WMS in the pipe is shown in Fig. 2. The WMS is designed as a matrix of 16×16 with a pitch of 2 mm between each stainless steel wire. The transmitter and the receiver are arranged 90° in a plane with a gap of 2 mm, that is, the thickness of the PCB is 2 mm. A total of 172 measuring points were distributed on the pipe cross-section. Fig. 3 is the schematic diagram of the WMS circuit, which is a dual-modality group circuit. The imaging frequency is up to 5 kHz. The noise / signal ratio is 0.2 %. Two groups of 128×128 matrices are supported at most. Matlab2019 Ra was used to process the data and analyze the gas-liquid concentration distribution at the longitudinal central section. In addition, the high-speed photography image processing method refers to Li Liu's paper[26].
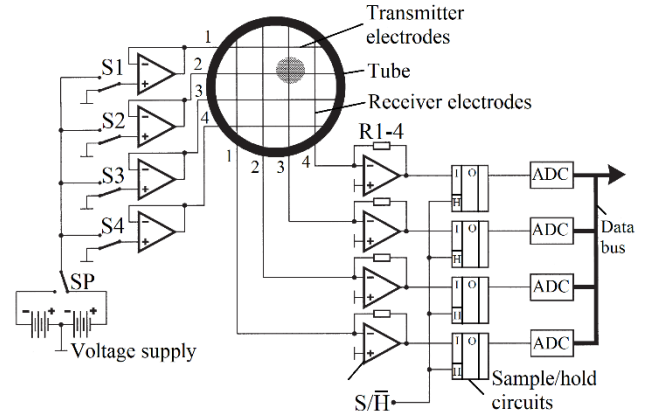
Fig. 3. Schematic diagram of WMS circuit unit.

III. RESULTS AND DISCUSSION

In the process of this study, a large number of data are recorded and analyzed to ensure the reliability of the research results. Next, the typical data will be presented.

A. The contrast analysis of the image of nonswirling gas-liquid flow patterns between WMS and high-speed photography
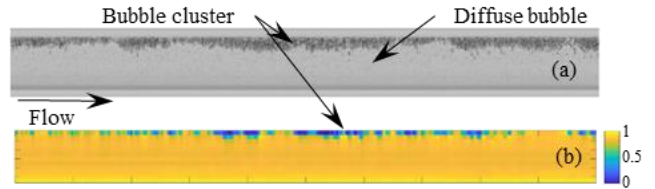
Fig. 4. Bubble flow (j_l = 1.88 m/s, j_g = 0.079 m/s).
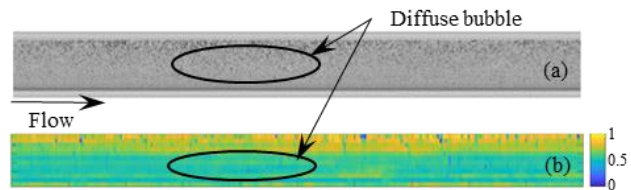
Fig. 5. Diffuse bubble flow (j_l = 2.27 m/s, j_g = 0.016 m/s).

Fig. 4 is a time-domain image of bubble flow. As can be seen from Fig. 4, high-speed photography (Fig. 4 (a)) can show more details than WMS (Fig. 4 (b)), including small dispersed bubbles in the liquid phase. WMS can only reflect the gas-liquid concentration distribution in the high-density bubble region. It is worth noting that WMS imaging (Fig. 5 (b)) has a large error in high-density diffused bubble flow. As can be seen from (Fig. 5 (a)), there are a lot of dispersed bubbles in the pipe, and the bubble concentration in the upper part of the pipe is slightly higher than that in the lower part of the pipe. However, in WMS images, it shows the opposite

results. This is because the conductivity of the gas-liquid mixture in the whole cross-section decreases at this time, while the bubble agglomeration phenomenon exists in the upper part of the pipeline, which makes the local conductivity drop obviously. The results show that the local gas concentration in the upper part of the pipeline section is too high and the liquid concentration in the upper part is generally too high.

Fig. 6 is a comparison of two imaging methods of stratified flow. It can be seen from Fig. 6 that under this condition, WMS and high-speed photographing results are highly consistent, which shows that under the condition of a clear phase interface, WMS recognition ability can be better.
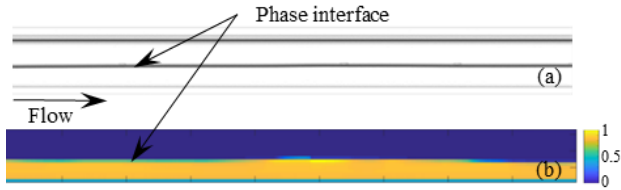
Fig. 6. Stratified flow (j_l = 1.18 m/s, j_g = 0.09 m/s).

Fig. 7 is a comparison of two methods in plug flow. It can be seen from Fig. 7 that WMS has a good ability to identify gas plug. In the region of liquid plug containing dispersed bubbles, the concentration of gas-phase in WMS imaging is lower than the real value (high-speed photography imaging). It also shows that the recognition ability of WMS is limited under the condition of microbubble dispersion. This is because the spacing between the screens is 2 mm. When the bubble diameter is less than 2 mm, WMS can't be identified by direct data. It can only reflect the distribution of bubbles in the liquid phase to a certain extent by setting a reasonable threshold value based on the super separation rate algorithm and data processing.
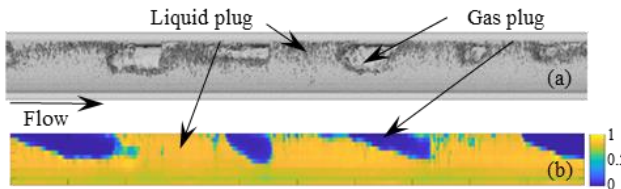
Fig. 7. Plug flow (j_l = 0.8 m/s, j_g = 0.393 m/s).

In the case of slug flow, the coincidence between WMS and high-speed photography is also good. However, due to the high gas velocity, the liquid interface fluctuates under the shear effect of the gas flow, which leads to the decrease of the definition of the phase interface at the gas slug region. As a whole, WMS has certain recognition ability for liquid slugs, gas slugs, pseudo slugs, and semi slugs in Fig. 8.
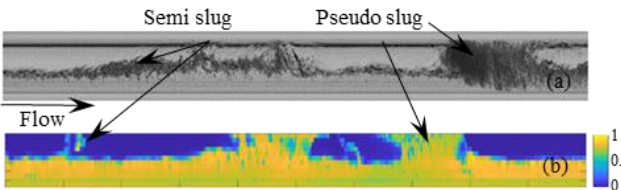
Fig. 8. Slug flow (j_l = 0.98 m/s, j_g = 3.2 m/s).

When the flow pattern develops to annular flow, as shown in Fig. 9, WMS has almost lost the ability of data acquisition

for the wall liquid film. In Fig. 9 (a), there is only a small amount of imaging data in the lower half of the pipeline with a thick liquid film. This shows that WMS can hardly measure thin liquid film.
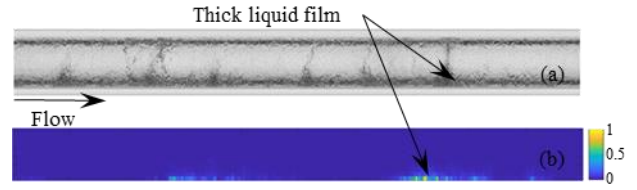
Fig. 9. Annular flow (j_l = 0.059 m/s, j_g = 20.17 m/s).

B. The contrast analysis of the image of swirl gas-liquid flow patterns between WMS and high-speed photography
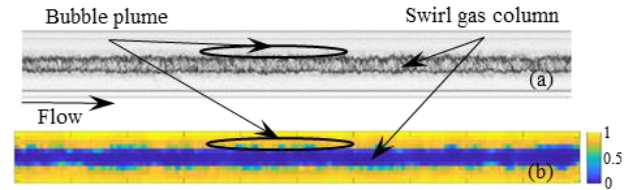
Fig. 10. Swirling gas column flow (j_l = 1.88 m/s, j_g = 0.079 m/s).

The swirling gas column flow is developed from the bubble flow under the action of the swirler [27]. As can be seen from Fig. 10 (a), there are a lot of tiny bubbles around the swirling gas core. These tiny bubbles can't be transferred to the center of the pipe by centripetal force alone to form the air core. In Fig. 10 (b), WMS imaging can better reflect this phenomenon. That is to say, there is a certain over a range between the swirling gas core and the swirling liquid annular. These transition regions reflect the existence of bubbles around the spiral core. In addition, when the diameter of the swirling gas core is small, the recognition ability of WMS to the swirling gas core decreases, and some areas with a small diameter cannot even be recognized effectively.
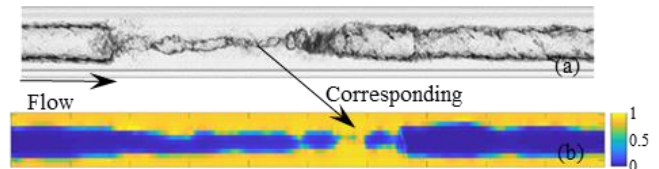
Fig. 11. Swirling intermittent flow (j_l = 0.204 m/s, j_g = 0.393 m/s).

The swirling intermittent flow is developed from plug flow or slug flow under the action of the swirler [27]. It can be seen from Fig. 11 that WMS can recognize the swirling intermittent flow well. However, the resolution of small diameter gas core in this flow pattern is not high, which is the same as that in small diameter swirling gas column flow.

The swirling annular flow developed from the non-swirling annular flow under the action of the swirler [27]. Like the annular flow, WMS cannot effectively distinguish the existence of the thin film.
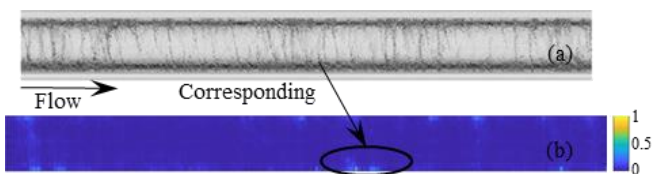
Fig. 12. Swirling annular flow (j_l = 0.059 m/s, j_g = 20.17 m/s).

Based on the swirling annular flow, continue to reduce the liquid flow and increase the gas flow to form the swirling ribbon flow, as shown in Fig. 13 (a). Because of the shearing force of the swirling airflow, the liquid film is sheared into a swirling liquid ribbon. In this process, a large number of small droplets are dispersed in the pipeline, which is captured and identified by the measuring points formed by the transmitter and the receiver when passing through the WMS, forming the image of Fig. 13 (b). Similarly, WMS cannot effectively identify the thickness of the swirling ribbon at this time. It can be seen that the identification ability of the WMS to the thin liquid film may need further optimization of the research work of the data post-processing algorithm, which cannot be effectively analyzed only by direct data.
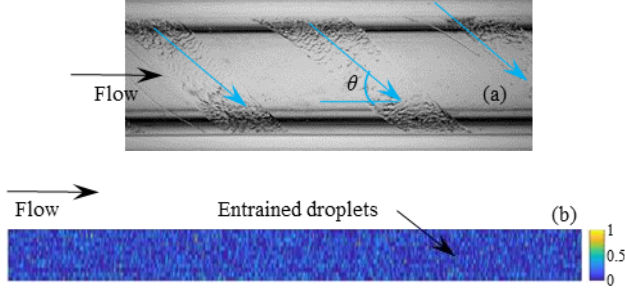
Fig. 13. Swirling ribbon flow (j_l = 0.02 m/s, j_g = 21.36 m/s).

C. Phase Interface Imaging

The analysis of the gas-liquid three-dimensional interface is helpful to better understand the phase distribution and flow pattern evolution in gas-liquid flow.
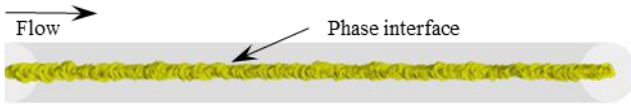
Fig. 14. The central gas core phase interface of swirl gas column flow (j_l=2.27 m/s, j_g=0.016 m/s).

The phase interface reconstruction of the swirling gas column flow is shown in Fig. 14. It can be seen from Fig. 14 that the shape of the central swirling gas core is irregular, but it shows that the equivalent diameter fluctuates continuously in a certain range along the flow direction. This is because the gas holdup is not completely consistent along the flow direction, but fluctuates in a certain range. Moreover, the instability of the swirling flow field will also aggravate the fluctuation of the diameter. On the whole, the reconstitution image of the swirling gas column phase interface can reflect the characteristics of this flow pattern.

Fig. 15 shows the interface imaging of the gas core in the center of the swirling intermittent flow. It is found that the diameter fluctuation of the central swirling gas column is very obvious. The characteristics of the axial section of the swirling intermittent flow in Fig. 11 are in good agreement. The gas plug or slug changes into a large diameter swirling gas core in the swirling field, and the tiny bubbles in the liquid plug slug gather to form a small diameter swirling gas core. The alternation of large diameter gas core and small diameter gas core leads to the large fluctuation. When the gas flow decreases, the intermittent fluctuation will weaken. In the process of phase interface imaging, some swirling gas cores with a smaller diameter cannot be identified, which is shown in the form of swirling gas core fracture in the phase interface image, at this time, the phase interface image is distorted. This

shows that the sensor in this paper is not suitable for the swirling intermittent flow with small gas flow, which may require further optimization of the data processing algorithm.
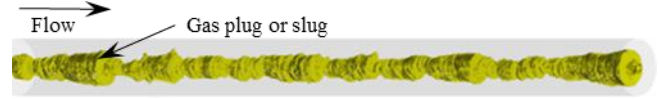
Fig. 15. The central gas core phase interface of swirl intermittent flow (j_l = 0.975 m/s, j_g = 3.2 m/s).

Fig.16 shows the reconstruction of the gas plug interface in plug flow. It can be seen from Fig. 16 that WMS has a good interface reconstruction ability for a large-scale gas plug. However, the resolution of the small gas plug or bubble dispersed in the liquid plug is lost.
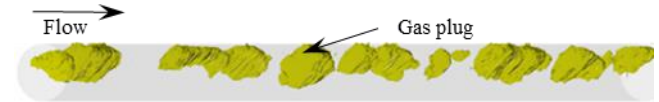
Fig. 16. The central gas core phase interface of plug flow (j_l = 0.8 m/s, j_g = 0.393 m/s).

(a) non-swirling annular film   (b) swirling annular film
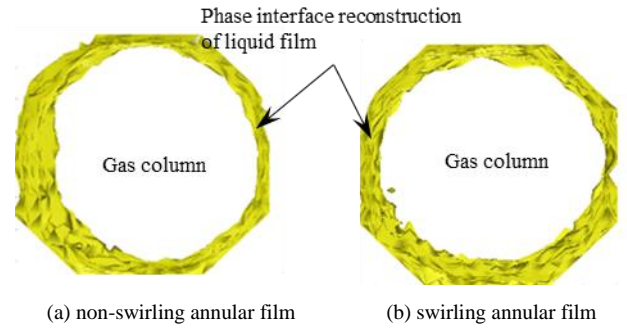
Fig. 17. The central gas core phase interface of plug flow (j_l = 0.059 m/s, j_g = 20.17 m/s).

Fig. 17 is the reconstruction image of the liquid film phase interface of the non-swirling and swirling annular flow. In Fig. 17 (a), it is found that the liquid film distribution of the non-swirling annular flow is not uniform. After passing through the swirler, under the shear effect of centrifugal force, the distribution area of the liquid film tends to be uniform. As shown in Fig. 17 (b), it can be seen that the distribution area of the liquid film is uniform under the shear force of tangential direction in the centrifugal force field. But whether it really reflects the thickness of the liquid film needs further research and verification. The phase distribution imaging of the longitudinal section in Fig. 9 and Fig. 12 can not reflect the existence of the liquid film, but it is displayed in the phase interface imaging, because only a series of measuring points are selected in the section imaging, and the phase interface imaging is the coupling of all measuring point data in the circumferential direction, and the imaging processing is carried out according to the data processing algorithm.

IV. CONCLUSIONS

In this paper, based on the high-speed imaging, the phase distribution imaging and phase interface imaging of typical flow patterns of non-swirling and swirling gas-liquid flow by WMS imaging technology are analyzed. It is found that the WMS has a good resolution for the phase distribution of bubble flow, plug flow, slug flow, swirling gas column flow and intermittent flow. The resolution of large bubble imaging

is better than that of small bubble imaging, and WMS has poor recognition ability for small bubble imaging. The ability of image recognition for the thin liquid film is also limited. In the future work, we will focus on the optimization of WMS design and data processing algorithms.

REFERENCES

[1] Kong R, Kim S. Characterization of horizontal air-water two-phase flow[J]. Nuclear Engineering and Design, 2016:S0029549316301765.

[2] Setyawan A, Indarto, Deendarlianto. Experimental investigations of the circumferential liquid film distribution of air-water annular two-phase flow in a horizontal pipe[J]. Experimental Thermal & Fluid Science, 2017, pp. 95-118.

[3] Dinaryanto O, Prayitno Y A K, Majid A I, Hudaya A Z, Nusirwan Y A, Widyaparaga A, et al. Experimental investigation on the initiation and flow development of gas-liquid slug two-phase flow in a horizontal pipe[J]. Experimental Thermal & Fluid Science, 2017, pp.93-108.

[4] Wrasse A D N, Vendruscolo T P, Santos E N D, Pipa D R, Moura H L, Castaldo F, et al. Capacitive Multielectrode Direct-Imaging Sensor for the Visualization of Two-Phase Flows[J]. Sensors Journal, IEEE, 2017, 17(24):8047-8058.

[5] Roitberg E, Shemer L, Barnea D. Measurements of cross-sectional instantaneous phase distribution in gas-liquid pipe flow[J]. experimental thermal & fluid science, 2007, pp.867-875.

[6] Ahmed W H . Experimental investigation of air-oil slug flow using capacitance probes, hot-film anemometer, and image processing[J]. International Journal of Multiphase Flow, 2011, pp.876-887.

[7] Hernández Cely Marlon M, Baptistella V E C, Rodriguez O M H . Study and characterization of gas-liquid slug flow in an annular duct, using high-speed video camera, wire-mesh sensor and PIV[J]. Experimental Thermal & Fluid Science, 2018, pp.563-575.

[8] Manfredo G, Beatrice B, Giorgio S, Colombo L P M. Image-based analysis of intermittent three-phase flow[J]. International Journal of Multiphase Flow, 2018:S030193221830171X-.

[9] Bofeng Bai, Liejin Guo, Liang Zhao. Research progress of on-line identification of gas-liquid two-phase flow pattern [J]. Progress in mechanics, 2001, pp. 437-446.

[10] Ito D, Prasser H M, Kikura H, Aritomi M. Uncertainty and intrusiveness of three-layer wire-mesh sensor[J]. Flow Measurement & Instrumentation, 2011, pp.249-256.

[11] Velasco P.A H F, Rodriguez O M H . Applications of wire-mesh sensors in multiphase flows[J]. Flow Measurement & Instrumentation, 2015, pp.255-273.

[12] Johnson I D. Method and apparatus for measuring water in crude oil: US 1987.

[13] Prasser H M, B Ttger A, Zschau J . A new electrode-mesh tomograph for gas-liquid flows[J]. Flow Measurement & Instrumentation, 1998, pp.111-119.

[14] Prasser H M, Beyer M, Boettger A, Carl H, Lucas D, Schaffrath A, et al. Influence of the pipe diameter on the structure of the gas-liquid interface in a vertical two-phase pipe flow[J]. Nuclear Technology, 2005, 152(1):p.3-22.

[15] Lucas D, Krepper E, Prasser H M. Development of co-current air-water flow in a vertical pipe[J]. International Journal of Multiphase Flow, 2005, pp.1304-1328.

[16] Silva M J D, Schleicher E, Hampel U. Capacitance wire-mesh sensor for fast measurement of phase fraction distributions[J]. Measurement Science & Technology, 2007, pp.2245-2251.

[17] Prasser H M, H Feli R . Signal response of wire-mesh sensors to an idealized bubbly flow[J]. Nuclear Engineering & Design, 2017:S0029549317301772.

[18] Kanai T, Furuya M, Arai T, Shirakawa K, Nishi Y. Three-dimensional phasic velocity determination methods with wire-mesh sensor[J]. International Journal of Multiphase Flow, 2012, pp.75-86.

[19] Prasser H M, Scholz D, Zippe C. Bubble size measurement using wire-mesh sensors[J]. Flow Measurement & Instrumentation, 2001, pp.299-312.

[20] Hoppe D, Grahn A, Peter Schütz. Determination of velocity and angular displacement of bubbly flows by means of wire-mesh sensors and correlation analysis[J]. Flow Measurement & Instrumentation, 2010, pp.48-53.

[21] Fuangworawong N, Kikura H, Aritomi M, Komeno T. Tomographic imaging of counter-current bubbly flow by wire mesh tomography[J]. Chemical Engineering Journal, 2007, pp.111-118.

[22] Weerin Wanjiraniran, Yuichi Motegi, Intrusive Effect of Wire Mesh Tomography on Gas-liquid Flow Measurement, Journal of Nuclear Science and Technology, 2005, pp.932-940

[23] Silva M J D, Thiele S, Abdulkareem L, Azzopardi B J, Hampel U. High-resolution gas–oil two-phase flow visualization with a capacitance wire-mesh sensor[J]. Flow Measurement & Instrumentation, 2010, pp.191-197.

[24] Da Silva M J, Dos Santos E N, Hampel U, Rodriguez, I H, Rodriguez O M H. Phase fraction distribution measurement of oil-water flow using a capacitance wire-mesh sensor[J]. Measurement Science & Technology, 2011, 22(10):104020.

[25] Eduardo N D S, Rigoberto E M M, Marco J D S. Advanced image processing of wire-mesh sensor data for two-phase flow investigation[J]. Latin America Transactions, IEEE (Revista IEEE America Latina), 2015, pp.2269-2277.

[26] Liu L, Wang K, Bai B. Experimental study on flow patterns and transition criteria for vertical swirling gas-liquid flow[J]. International Journal of Multiphase Flow, 2019, 122:103-113.

Flow Visualization of Centrifugal Pump by the Combination of LIF and PIV

Qi Zhang*
School of nuclear science and engineering
Shanghai Jiao Tong University
Shanghai, China

Hanyang Gu
School of nuclear science and engineering
Shanghai Jiao Tong University
Shanghai , China

Shuai Liu
School of nuclear science and engineering
Shanghai Jiao Tong University
Shanghai , China

Junlong Li
School of nuclear science and engineering
Shanghai Jiao Tong University
Shanghai , China

Sichao Tan
Fundamental Science on Nuclear Safety and Simulation Technology Laboratory
Harbin Engineering University,
Harbin, China

Jianke Su
Support Center of Nuclear Technology, CAEA
Beijing, China

Abstract—**The visual experimental system for the centrifugal pump was established. The mock centrifugal pump with six curving semi-open impellers was manufactured by polymethyl methacrylate. The flow characteristics inside the mock centrifugal pump was studied by the combination of laser induced fluorescence (LIF) and particle image velocimetry (PIV). The polyamide resin particles were dissolved into the fluid of the circulation for PIV analysis, and fluorescent dye rhodamine B was head-on injected to the impellers for LIF analysis. The front plane of the impellers was covered by Nd: YAG laser sheet, one CCD camera collected the reflected laser light from the particles. Synchronously, another camera collected the fluorescence of rhodamine B. The velocity distribution characteristic was presented, and the time evolution of injected dye in the channel between impeller and septum was obtained. The inlet fluid moved circumferentially along the impeller under centrifugal force, the central region had smaller velocity. The injected dye gathered into blocks near the central area, and was evenly distributed near the edge of the pump impellers. The combination of LIF and PIV was proved to be credible method in the research of centrifugal pump.**

Keywords—centrifugal pump; flow visualization; LIF; PIV

I. INTRODUCTION

The centrifugal pump is a common and essential device in the engineering field. The centrifugal pump is generally associated with issues as cavitation and flow induced vibration. The internal flow characteristics of centrifugal is pivotal to the reliability and safety. Acosta[1] measured the velocity and pressure distribution in the centrifugal pump by pressure gage. Acosta found the internal energy loss was concentrated near the suction surface of the impeller. Lennemann[2] and Howard[3] found that shrouded impeller had a higher fluid leakage than the unshrouded impeller. The fluid leakage was restrained by installing a cover plate in the front of impellers. The above researches used contact detectors such as hot film probe, the original flow field was disturbed by these detectors. The installation of the experimental apparatus was also very difficult.

The visualization technology has emerged to study the hydraulics in complex object. Wang[4] and Qi[5] verified that LIF and PIV have high sensitivity and measurement accuracy, and LIF and PIV are applicable for the narrow channel. In this

paper, the LIF and PIV were combined to study the flow and mixing characteristics of a semi-open centrifugal pump with six curving impellers. This paper illustrates the operation of the combination of LIF and PIV. The particles were evenly mixed into the fluid of circulation, the velocity distribution in front of the impellers was obtained. The fluorescent dye was injected into the pump as tracer, the tracer transportation process was shown. The vortex motion and the mixing characteristics in the centrifugal pump were analyzed.

II. PRINCIPLE OF LIF AND PIV

Both LIF and PIV are performed on the basis of laser generator, which has a stable optical wavelength and high energy intensity. The object of LIF technology is the emission fluorescence from fluorescent dye, and the object of PIV is the reflected light of the particle. Fig.1. shows the de-excitation process of dye molecule, the fluorescence is emitted by stokes redshift effect[6]. The intensity of fluorescence is as follows:

$$I = K_{opt}V_cI_0\phi\varepsilon_1Ce^{-C(\varepsilon_1b+\varepsilon_2e)} \tag{1}$$
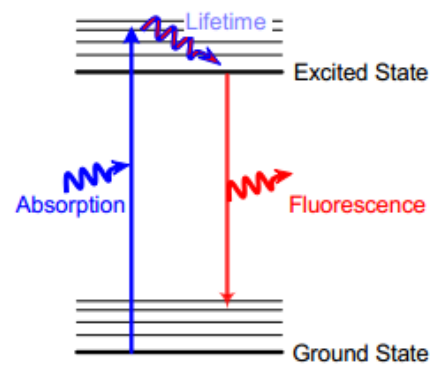
Fig.1. The schematic diagram of LIF

The term $e^{-C(\varepsilon_1b+\varepsilon_2e)}$ is decided by optical path, and is nearly unchanged with low concentration. In this case, the fluorescent intensity I positively increases with dye concentration. CCD camera can record the fluorescence in the form of dimensionless gray value.

Fig.2. shows the schematic diagram of PIV[7]. The particles with optical scattering property is diffused into the experimental fluid[8]. The incident laser is reflected by the particle. The transient location of the particles is captured by

CCD camera. The basic observation window was composed by 4 × 4 pixels, there were 2～4 particles in the observation window. With a short period Δt (the inverse of camera frame), the transverse and longitudinal displacement distance of the particle is analyzed, the velocity distribution is constructed by the particle track. The measuring error of PIV is decided by particle trajectory, mis-matching error and physical distance of the devices[8].
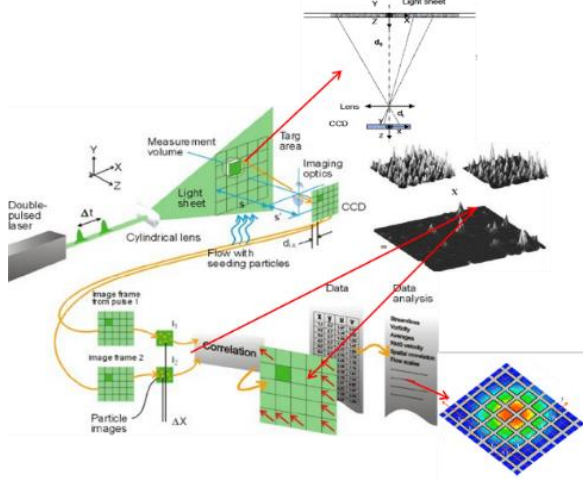
Fig.2. The schematic diagram of PIV
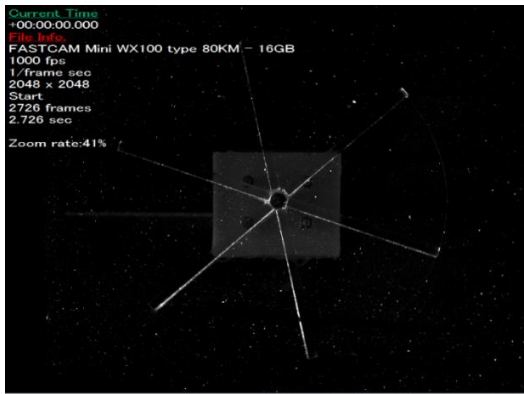
Fig.3. The particle images inside the centrifugal pump

III. THE EXPERIMENTAL SET UP

Fig. 4 shows the schematic diagram of experimental circuit. Deionized water in the water tank was driven by the experimental pump, and an auxiliary pump in series on the circulation to increase the driving force. The flowrate was controlled by the bypass valve, the water temperature and operation pressure were monitored by thermocouple and pressure gage. Fig. 5 shows the test section of the centrifugal pump. The main body was an industrial centrifugal pump, the stainless shell was removed and was replaced by the rectangular transparent shell. There were four exhaust vents on the top of the rectangular shell, the air in the shell was removed before formal experiment. A stainless pipe with 1mm inner diameter was inserted into the test section. The fluorescent dye was injected to the front of the impeller by a peristaltic pump. Fig. 6 shows the layout of test section and the main devices. The fan-shaped laser sheet was generated by the laser generator (the wavelength was 532 nm), and it located 5mm in front of the six wings impeller. The particle was polyamide resin, and the fluorescent dye was Rhodamine B. The emergent light from test section was divided into two shares by a spectroscope, and was respectively recorded by two CCD

cameras. The shutting speed of the CCD cameras was 1000 frame.
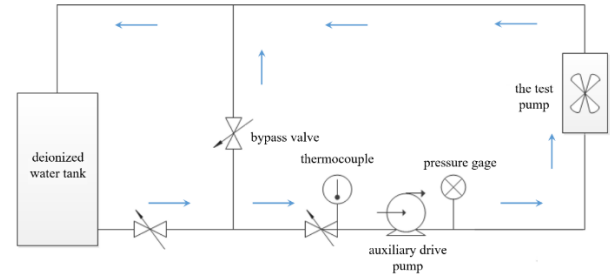
Fig. 4. The schematic diagram of experimental circuit
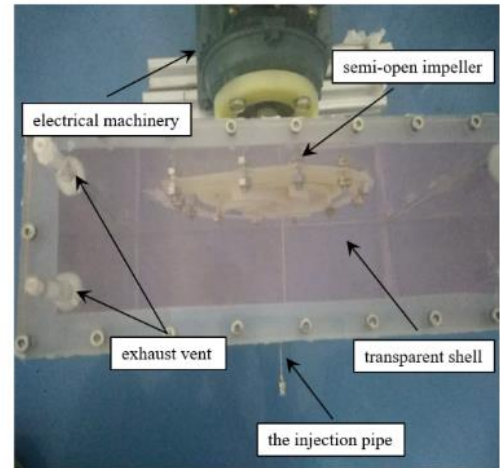
Fig. 5. The test section of the centrifugal pump
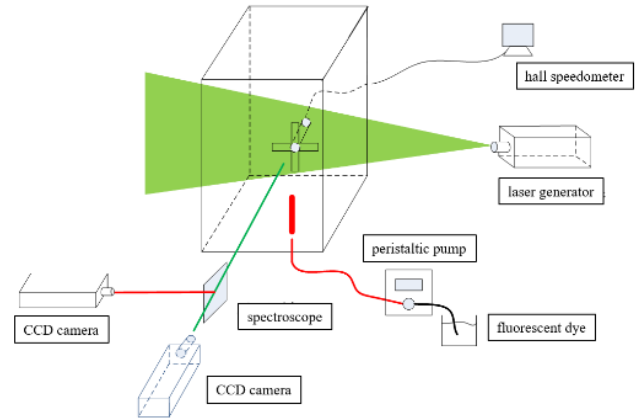
a. schematic of experimental set up
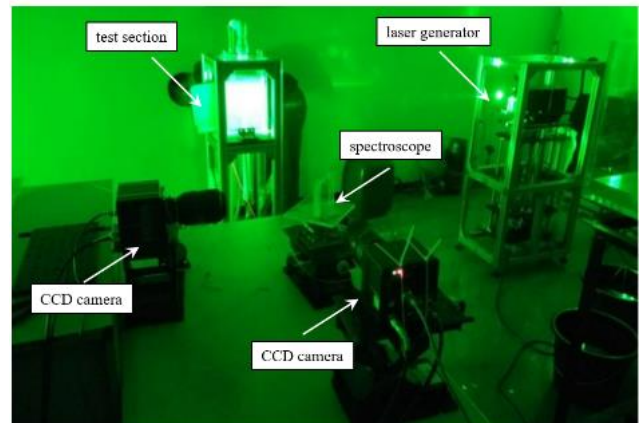
b. physical diagram of experimental set up
Fig. 6. The layout of test section and main devices

430

As described in the experimental principle, the object of LIF is fluorescence while PIV concerns the reflected laser light. The part of the laser energy was dissipated into the experimental fluid, thus the fluorescence has weaker intensity and higher wavelength. The intensity of fluorescence was mainly decided by dye concentration and the fluid temperature. The operation power of laser generator was 10W, the fluid temperature variation was ignored, the fluid temperature was 25 °C. The fluorescence and reflected laser were overlapped, they interfered each other to some extent, the dye concentration of 10∼100mg/L was respectively validated (see Table). The injection speed of peristaltic pump was 60mL/min.

Table I The Experimental Ccondition of Centrifugal Experiment

No.	rotate speed （r/min）	dye concentration （mg/L）	dye injection speed (mL/min)
1	180	100	60
2	180	50	60
3	180	25	60
4	180	10	60

IV. Results and Discuss

Fig.7 shows the images of fluorescence and reflected laser under different concentration. The motor shaft and gasket appeared in the background of the image. Zhang[9] introduced the method of background noise elimination of LIF. With low dye concentration (*C*=10 mg/L), the gray value of dye tracer was slight and distinguished. When the dye concentration increased to 50mg/L, the track of particle was totally covered by the injected dye concentration, which obviously affected the particle recognition. Both LIF and PIV were availably when the dye concentration was 25mg/L. For PIV data processing, the dye tracer was eliminated by noise identification. For LIF test, there was a filter on the corresponding camera. This filter reflected the laser light and the fluorescence could penetrate the filter.

(a) 10mg/L (b) 25mg/L

(c) 50mg/L (d) 100mg/L
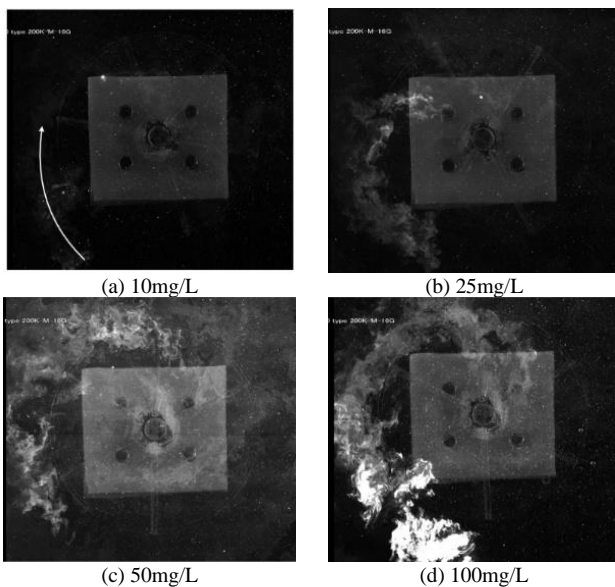
Fig.7. The image of fluorescence and reflected laser under different concentration

Fig. 8 shows the original tracer distribution. The location of dye tracer inlet was marked, the injected dye spread along clockwise rotation of impellers. The fluid in the pump deviated from the central under the action of centrifugal force. At the time of 0.48s, the injected dye distributed in discrete blocks, and the block size decreased as it deviated the injected

pipe. With constant injecting, these dye blocks started to piece together near the end of impellers (T=1.78s). At the time of 5.68s, the blocks still existed near the injected pipe, and the previous injected dye uniformly distributed in the lower left corner.

(a) 0 s (b) 0.48 s

(c) 1.78 s (d) 5.68 s
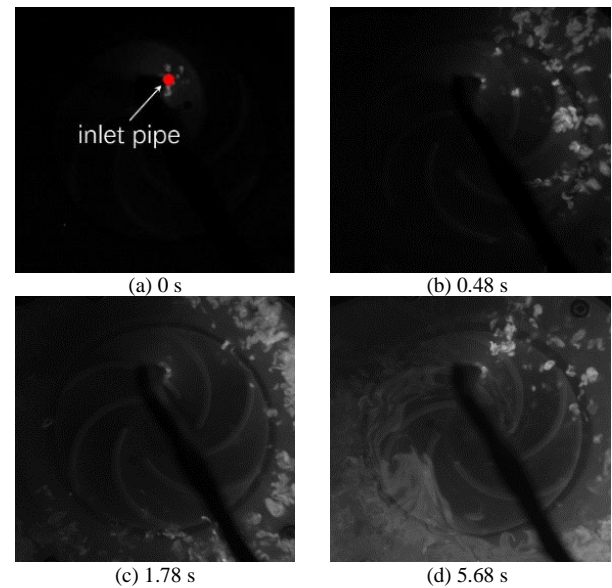
Fig. 8. The time evolution of injected dye tracer

(a) 0 s (b) 0.48 s

(c) 1.78 s (d) 5.68 s
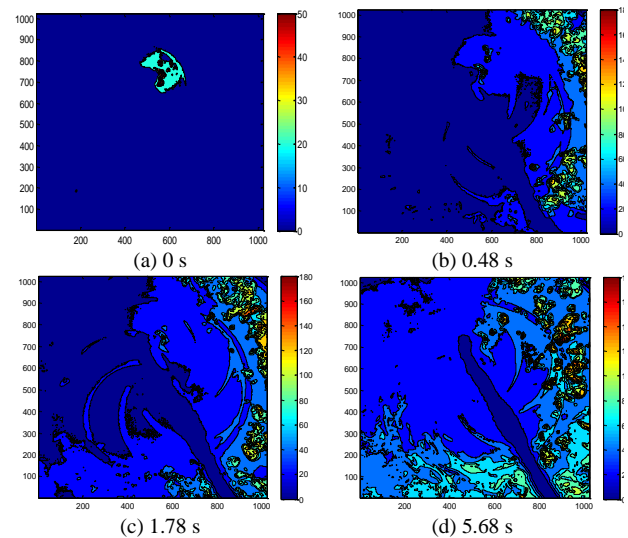
Fig. 9. The time evolution of dye concentration

Fig. 9 shows the processed concentration images, which were more intuitive and phenomenal than the original gray value image. From Fig. 9 (b) and (c) the upper part of the transparent pump was occupied by fluorescent dye, the concentration was low and uniform distributed. At the time of 5.68s, the whole container of the pump was filled with low concentration of dye; the subsequent injected dye spread in the same form as T=0.48s.

Fig. 10 shows velocity distribution of PIV technology. The velocity vector shows that the fluid forms swirling flow clockwise. The mass flowrate of injected dye was negligible to the bulk flow and the flow field was not affected. From velocity distribution cloud map, the fluid was stagnant near the central of the pump central. The central part of the impeller has no blade, and the centrifugal force was negligible near the central. Adler[10] reported similar phenomenon with LDV technology. The near the front of the impeller, the velocity is obviously larger. The injected dye

spread rapidly by the stronger swirl could not stuck in these area. The velocity was smaller and gentle in area far from the central, which corresponded to the stable concentration distribution of Fig. 9 (c) and (d).

Fig. 11 is the composite image of dye distribution and velocity vector. The optical data of LIF and PIV was simultaneously shot. Only a limited tracer entered the central of the pump by the turbulent diffusion. The tracer blocks have significantly reduced velocity. For the area which has small velocity gradient, the tracer was also evenly distributed.
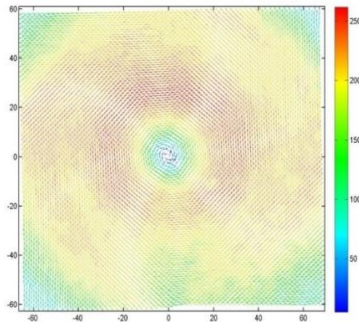
Fig. 10. The velocity distribution of PIV technology
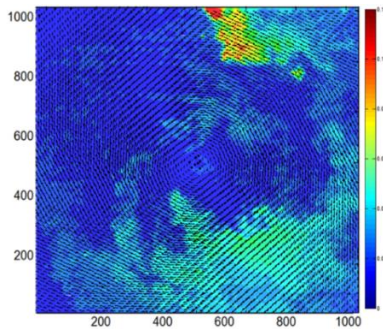
Fig. 11. The synchronous image of LIF and PIV

V. CONCLUSION

In this paper, the combination of LIF and PIV was trialed. The flow visualization in a transparent centrifugal pump was conducted. The emergent light of test section was divided into two shares. Two CCD cameras operated simultaneously, one share of the light for LIF experiment and another for PIV experiment. The commended fluorescent dye concentration was 25mg/L in this experiment.

The mixing characteristics and velocity distribution of the centrifugal were simultaneously obtained. The tracer was injected in front of the impeller, the time evolution of tracer was obtained. The tracer gathered into blocks near the inlet pipe, and uniformly distributed in the opposite area. The synchronous velocity distribution was obtained by PIV. The rotational flow was formed near the central area, and clockwise spread towards the outer ring area.

REFERENCES

[1]   Acosta, A.J. and R.J.T.o.t.A. Bowerman, *An experimental study of centrifugal-pump impellers.* 1957. **79**(4): p. 1821-1839.
[2]   Lennemann, E. and J. Howard, *Unsteady flow phenomena in rotating centrifugal impeller passages.* 1970.
[3]   Howard, J. and C. Kittmer, *Measured passage velocities in a radial impeller with shrouded and unshrouded configurations.* 1975.
[4]   Wang, X., et al., *Flow visualization and mixing quantification in a rod bundle using laser induced fluorescence.* 2016. **305**: p. 1-8.
[5]   Qi, P., et al., *Experimental investigation of the turbulent flow in a rod bundle channel with spacer grids.* 2019. **130**: p. 142-156.
[6]   Lemoine, F., M. Wolff, and M.J.E.i.F. Lebouche, *Simultaneous concentration and velocity measurements using combined laser-induced fluorescence and laser Doppler velocimetry: Application to turbulent transport.* 1996. **20**(5): p. 319-327.
[7]   Raffel, M., et al., *Particle image velocimetry: a practical guide*. 2018: Springer.
[8]   Nishio, S.J.I.R.P. and Guidelines, *Uncertainty Analysis: Particle Imaging Velocimetry (PIV).* 2008.
[9]   Zhang, Q., et al., *Experimental study on mixing phenomenon inside reactor down-comer under single-loop injection using laser induced fluorescence.* 2019. **117**: p. 103046.
[10] Adler, D. and Y.J.J.o.M.E.S. Levy, *A laser-Doppler investigation of the flow inside a backswept, closed, centrifugal impeller.* 1979. **21**(1): p. 1-6.

.

Sensitive Detection of Trace Pesticide Residue Implemented by Terahertz All-Dielectric Metamaterial

Zijian Cui
Key Laboratory of Ultrafast Photoelectric Technology and Terahertz Science in Shaanxi
Xi'an University of Technology
Xi'an, China
cuizijian_harbin@163.com

Yue Wang*
Key Laboratory of Ultrafast Photoelectric Technology and Terahertz Science in Shaanxi
Xi'an University of Technology
Xi'an, China
wangyue2017@xaut.edu.cn

Xiaoju Zhang
Key Laboratory of Ultrafast Photoelectric Technology and Terahertz Science in Shaanxi
Xi'an University of Technology
Xi'an, China
zxj11303@126.com

Lisha Yue
Key Laboratory of Ultrafast Photoelectric Technology and Terahertz Science in Shaanxi
Xi'an University of Technology
Xi'an, China
YueLisa_Y@163.com

Xinmei Wang
School of Automation and Information Engineering
Xi'an University of Technology
Xi'an, China
wangxinmei@xaut.edu.cn

Chen Yang
School of Automation and Information Engineering
Xi'an University of Technology
Xi'an, China
yangchen@xaut.edu.cnline

Hui Hu
Key Laboratory of Ultrafast Photoelectric Technology and Terahertz Science in Shaanxi
Xi'an University of Technology
Xi'an, China
fourneyron@126.com

Kuang Zhang
Department of Microwave Engineering
Harbin Institute of Technology
Harbin, China
zhangkuang@hit.edu.cn

Yue Wang*
Key Laboratory of Engineering Dielectrics and Its Application Ministry of Education
Harbin University of Science and Technology
Harbin, China
wangyue2017@xaut.edu.cn

Abstract—**A novel all-dielectric terahertz metamaterial is proposed, which has a broadband absorbance spectrum comprising two resonance peaks. The experiment results indicate that the designed metamaterial has great application potential as a benomyl sensor. Also, the regression analysis demonstrated the great linear correlation between the absorbance intensity and the concentration of benomyl. The correlation coefficient reaches 0.9954 and sensitivity of 0.0017 ppm^{-1}. These results indicate that the all-dielectric terahertz metamaterials can be used in the highly sensitive and stable detection of trace pesticide residues, unlocking a host of feasibility for future applications in the fields of biosensing and substances detection.**

Keywords—terahertz, metamaterial, absorbance, benomyl, sensor

I. INTRODUCTION

Rapid and accurate detection of pesticide residues has drawn great attention to the inspection of agricultural products [1]. However, these substances are toxic and dangerous [2]. Still and all, the use of pesticides is inevitable because of the purpose of increasing the yield of agricultural products [3]. The pesticides are widely used in vegetables, fruits, meats production, and so on. Long-term accumulation of pesticide residues will have a certain impact on human health [4]. At present, trace pesticide analysis and detection technologies are divided into immunoassay [5,6], gas chromatography [7], and chromatography mass spectrometry technology [8-10]. These techniques are sensitive and reproducible, yet commonly unneglectable expensiveness and time-consuming requiring of sample pretreatment operations are still problem. Therefore, it is very important to develop an advanced, effective, rapid, and non-destructive detection technology for pesticide residues.

As the terahertz technology advanced, especially the emergence of reliable radiation sources, terahertz waves have attracted extensive research attention because of its low photon energy and good penetration, which benefits non-
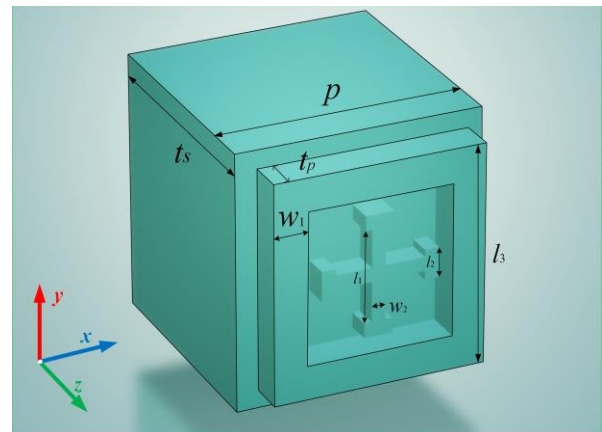
Fig. 1. Schematic diagram of metamaterial unit cell

destructive detection [11,12]. In addition, since the terahertz band contains the vibration and rotational frequencies of most biomolecules, making terahertz waves a promising candidates of biosensing applications. However, most substances respond weakly in the terahertz band, and many dielectric materials exhibit transparent properties [13-15]. This makes it difficult

---

to detect trace pesticide residues. Currently, the tablet analysis method is still the main research methods for pesticide residues in the terahertz frequency [16-18]. The fingerprint line of pesticides can be detected by this method, nevertheless, the sensitivity of the detection is still low, and only samples of milligrams scale can be detected.

Metamaterials are electromagnetic materials based on artificial periodical structure. Due to its unique electromagnetic properties, metamaterials exhibit differences from those of traditional natural materials, which has attracted extensive research attention and shown great application potential in the fields of cloaking [19], imaging [20], perfect lens, and sensing [21,22]. Their resonance frequencies of metamaterial depend on the artificial design, and the local field at the resonance frequency can be enhanced [23]. These properties make it possible to amplify the signal of the substance by using a terahertz metamaterial. Detection by using terahertz metamaterials is gradually arising in cancer cell detection, refractive index detection, and other fields.

Here, we propose a high sensitivity method for the detection of benomyl using a novel all-dielectric terahertz metamaterial device. The absorbance spectrums of the designed metamaterials treated by benomyl with various concentrations are measured using the terahertz time-domain system. Furthermore, the relationship between concentration and amplitude is analyzed. The absorbance intensity shows a linear relationship with trace benomyl concentration, which demonstrates that the all-dielectric metamaterial can be used as a fast, sensitive, and accurate sensing devices for the determination of trace pesticides, provides a variety of application feasibilities in biochemical sciences.

II. STRUCTURE DESIGN AND METHOD

A. The Structure Design of Metamaterial

The patterns of metamaterial unit cells are etched into a heavily–doped n-type silicon substrate. The pattern consists of an external square loop and an internal concentric nesting of the Jerusalem cross structure, as illustrated in Figure 1. The thicknesses of the silicon structure and the patterns are $t_s = 300$ μm and $t_p = 50$ μm, respectively. The dimensions of the centric Jerusalem cross are set as follows: $l_1 = 80$ μm, $l_2 = 30$ μm, and linewidth $w_2 = 15$ μm. The width of the exterior square loop is $w_1 = 40$ μm and the length is $l_3 = 220$ μm. The lattice constant is $p = 260$ μm. The repeated arrangement of the unit cells around constitutes the overall structure of the metamaterial.
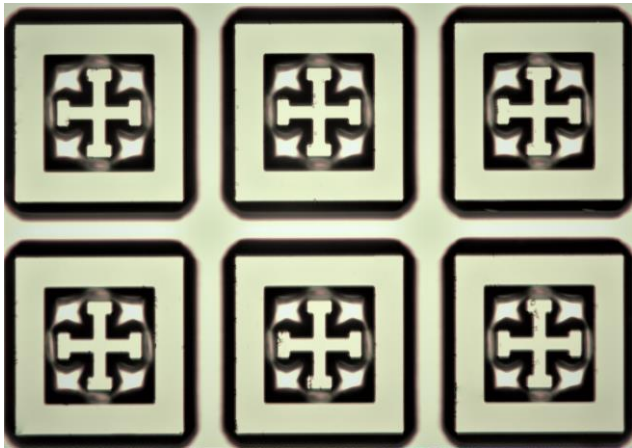
Fig. 2. Optical photomicrograph of metamaterial sample
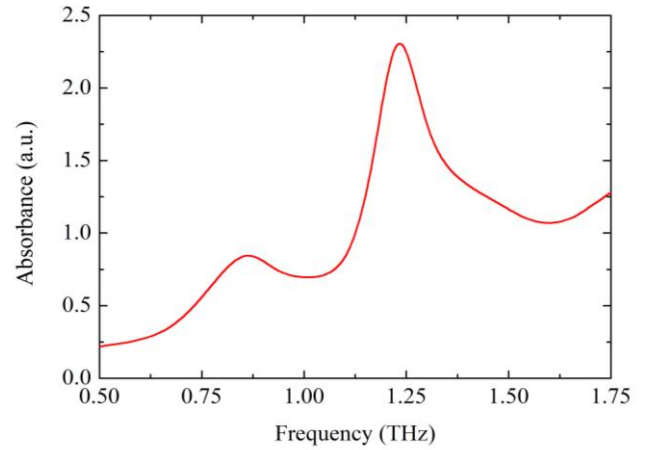
Fig. 3. The absorbance spectrum of metamaterial sample, from the average results of ten measurements.

By using conventional photolithography and deep reactive-ion etching based on large-scale microfabrication techniques, this all-dielectric metamaterials can be fabricated. The optical photomicrograph of the fabricated metamaterial is shown in Fig. 2.

B. Preparation of Pesticide

Analysis grade benomyl standard material (≥99.0%) was homogenized in an agate mortar, screened for 100 mesh. And then, the screened benomyl is dissolved in petroleum ether. First, 5 mg benomyl standard solid powder is mixed with 50 mL petroleum ether to prepare the pesticide solution sample
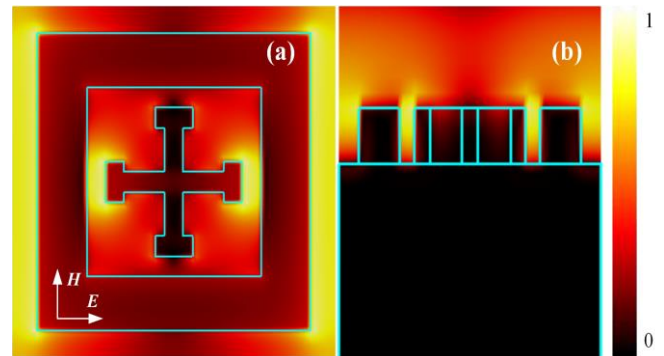
Fig. 4. Simulated electric field distribution. (a) From x-o-y plane; (b) from y-o-z plane.

with a concentration of 100 ppm. Then, by dilution with 100 ppm benomyl solution, benomyl solutions with concentrations of 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, and 52 ppm were prepared. Finally, the diluted benomyl solutions were evenly mixed by centrifugal oscillation.

C. Set of Terahertz Time-domain Spectrometer

The reflection test of metamaterial samples treated with benomyl solution of different concentrations was carried out using CCT-1800 terahertz time-domain spectrometer (THz-TDS). CCT-1800 has an effective frequency range of 0.1–4 THz with 30 GHz spectral resolution, the spot diameter is 3 mm, and the signal-to-noise ratio (SNR) at 0.5 THz is 70 dB. Nitrogen is continuously infused into the THz-TDS system to prevent the influence of the moisture during the entire experiment. The experiment environment is maintained at
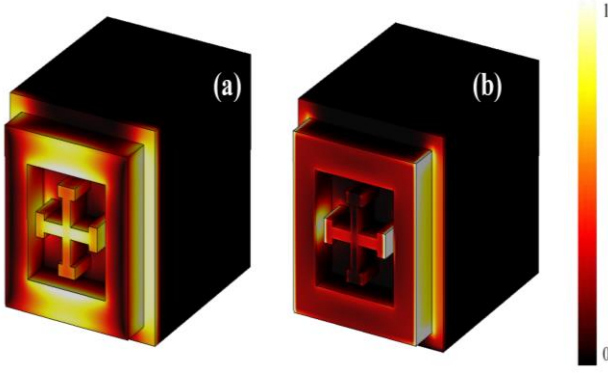
Fig.5. Simulation of magnetic field distribution and power flow distribution.

room temperature (27°C) and humidity less than 5%. Time-domain spectra of golden mirror reference and metamaterial samples are measured, and then, transform into frequency-domain using the fast Fourier transform.

III. RESULTS AND DISCUSSION

A. Measurements of Metamaterial

The metamaterial sample is examined experimentally via reflection-mode. Because of the heavy doping concentration, terahertz waves can hardly penetrate the designed device. Therefore, the absorption spectrum of the device can be obtained by reflection test. THz wave emitted from a photoconductive antenna is reflected from both the metamaterial sample and a golden mirror reference. The incident waves are at an angle of 30°. The metamaterial are measured in the case of TE polarizations, and the measured absorbance is shown in Figure 3. The results are based on an average of ten-time measurements. It can be seen from the absorbance spectrum that the designed metamaterial device achieves broadband absorption characteristics and has two resonance peaks at around 0.89 THz and 1.25 THz. This resonance peaks provide the ability to detect trace pesticide residues, which will be discussed in more detail later.
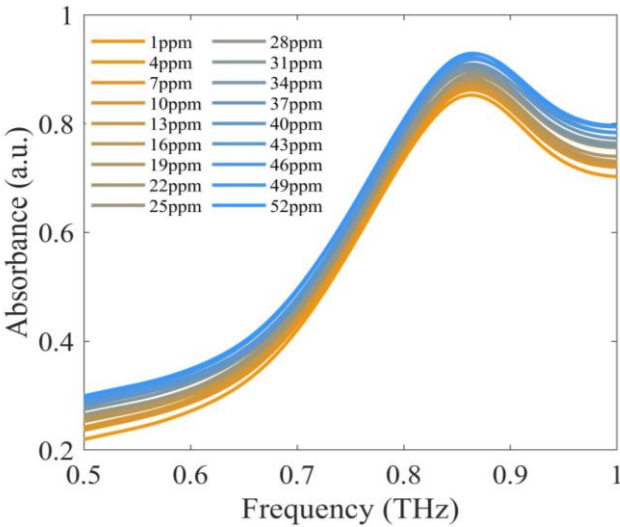
B. Simulations of Metamaterial

In order to get an in-depth insight into the physical mechanism of absorption, we use CST Microwave Studio to simulate the field distribution at the 0.89 THz positions. The incident electric field is along the *x*-direction, the boundary condition is set as the perfect electric boundary condition in the *x*-direction, and the *y*-direction is the perfect magnetic boundary condition, to mimic the periodic distribution. The dielectric properties of the silicon are described by the Drude model in the simulation [24].

$$\varepsilon = \varepsilon_\infty - \omega_p^2/(\omega(\omega + i\Gamma)) \qquad (1)$$

where ω_p is the plasma frequency, Γ is Drude collision frequency. $\varepsilon_\infty = 11.7$, and ε_0 is the permittivity of vacuum.

As illustrated in Fig.4, the electric fields are concentrated on the gap produced by the external square loop and Jerusalem cross structure. Where, Fig. 4 (a) and Figure 4 (b) are the results of the *x-o-y* plane and *y-o-z* plane, respectively. In particular, the electric fields clustered strongly at the gap between the Jerusalem cross and external square loop, and also, the gap between square loops of different units.

Similar to the confinement of the electric field, the magnetic field also accumulates at 0.88 THz, as shown in Fig.5 (a). It can be seen from the distribution of the overall structure that the magnetic field is concentrated on the lever of the external square loop and Jerusalem cross structure which are in the direction parallel to the incident electric field. In addition, the magnetic field also accumulates at the gap between square loops of different units.

The power flow distribution is shown in Fig.5 (b), which explains the occurrence of the absorption. Due to the aggregation effect of the metamaterial structure along the direction of the incident electric field, this local enhancement effect makes most of the energy dissipate in the gaps of the structure. However, the local field enhancement makes it possible to amplify the effects of pesticides on absorption characteristics, which provides the possibility to detect trace pesticide residue.

Fig. 6. Absorption spectra of metamaterials treated with benomyl solution of different concentrations
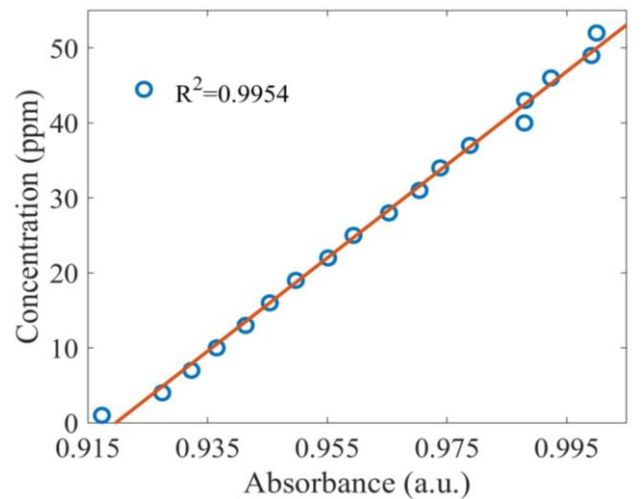
Fig. 7. The regression curves established based on the spectral peaks intensity

C. Enhanced Benomyl Sensing

The ability for qualitative and quantitative biological sensing of the proposed all-dielectric metamaterial structure is demonstrated in the present work. We demonstrate that this all-dielectric metamaterial based on heavily-doped silicon can be used as a fast, stable, highly sensitive, and accurate benomyl sensor in the terahertz band. In the experiment, the benomyl solution is dripped uniformly on the surface of the metamaterial sample using a pipette. Each time, 10 μL of benomyl solution completely covers the surface of the device. Then, the metamaterial is placed in a dust-free environment until the solution is air-dried. Finally, the reflection spectrums of the treated metamaterial are measured. The test is repeated 10 times at each concentration and the spectrum results are averaged. The SNR of the spectrum peak at 0.89 THz is much higher than that at 1.25 THz. Therefore, the absorption peak of 0.89 THz is chosen to explore the enhanced benomyl sensing phenomenon of all-dielectric metamaterials.

We tested the absorbance spectra of metamaterials treated with benomyl solution at concentrations of 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, and 52 ppm. Figure 6 shows the normalized absorbance spectra at different concentrations from 0.5 THz to 1.0 THz for the sake of intuition. As illustrated in Figure 6, the intensity of the absorption peak increases continuously with the increasing concentration of benomyl solution. Also, the intensity of the absorption peak of the metamaterial sample, which is not treated with benomyl solution, is lower than the treated. After being air-dried on the surface of the metamaterial, benomyl enters into the designed structures and changes the dielectric environment in the gap. The above analysis of the field distribution shows that at the resonant frequency, the electric and magnetic fields are strongly restricted in the gaps between the structures and the unit cells, which amplifies the small dielectric variation in the gaps on the absorbance spectrum of the metamaterial. Thus, the absorbance intensity changes with the variation of benomyl solution concentration.

Furthermore, the relation between the concentration of benomyl solution and the absorbance intensity is analyzed. The relation of the metamaterial samples treated with benomyl solution is obtained by extracting the peak intensity of the metamaterial samples with different concentrations at around 0.89 THz, as shown in Figure 7. The blue circles in Figure 7 represent the data tested in the experiment, and the orange line is the result obtained by the regression analysis of the experimental data. As can be seen from Fig.7, the experimental data are clustered on both sides of the regression line, and the correlation coefficient reaches 0.9954. This indicates that there is a good linear correlation between the concentration of benomyl solution and the absorbance intensity. Therefore, the intensity value can be used as the key index to detect the concentration of benomyl. Besides, the sensitivity of the designed metamaterial to detect the concentration of benomyl is 0.0017 ppm^{-1}, and the minimum detectable concentration is 1 ppm. However, as the dielectric environment changes, the formant should experience a frequency shift that does not show up in our experiments. For the structure of metamaterials in this paper, the influence on amplitude is much greater than that on frequency. The frequency shift may be smaller than the spectral resolution. To improve the detection sensitivity, methods such as improving the quality factor of the absorption peak and designing the resonance frequency to be consistent with the characteristic peak of the pesticide can be adopted. This work demonstrates the ability of all-dielectric metamaterials to trace pesticide residues detection sensors and provides a fast and efficient detection method.

IV. CONCLUSION

In the present work, an all-dielectric terahertz metamaterial based on heavily-doped silicon is proposed. The metamaterials are designed to demonstrate their potential as sensors for benomyl residues. The results showed that there is a good linear relationship between the concentrations of benomyl solutions and the absorbance intensity at around 0.89 THz. It is of great significance to realize a precise terahertz biosensor that is cheap, stable, fast, efficient, and highly sensitive.

REFERENCES

[1] G. Zhao, H. Wang, and G. Liu, "Advances in biosensor-based instruments for pesticide residues rapid detection," Int. J. Electrochem. Sci., vol. 10, pp. 9790-9807, December 2015.

[2] R. Mesnage, G. Seralini, "Editorial: toxicity of pesticides on health and environment," Front. Public Health, vol. 6, pp. 268, September 2018.

[3] Z. Zhong, B. Aotegen, H. Xu, and S. Zhao, "Structure and antimicrobial activities of benomyl phenyl-thiosemicarbazone-chitosans," Int. J. Biol. Macromol., vol. 50, pp. 1169-1174, May 2012.

[4] M. Ye, J. Beach, J. Martion, and A. Senthilselvan, "Pesticide exposures and respiratory health in general populations," J. Environ. Sci., vol. 51, pp. 361-370, January 2017.

[5] H. Jiang, M. Fan, "Multi-analyte immunoassay for pesticides: a review," Anal. Lett., vol. 45, pp. 1347-1364, 2012.

[6] Y. Liao, X. Cui, G. Chen, Y. Wang, G. Qin, M. Li, et. al, "Simple and sensitive detection of triazophos pesticide by using quantum dots nanobeads based on immunoassay," Food Agric. Immunol., vol. 30, pp. 522-532, January 2019.

[7] E. Ueno, "Application of gas chromatography/mass spectrometry for residue analysis of pesticides residue analysis of pesticides in foods using GC-MS and GC-MS/MS (Part I)," J. Pestic. Sci., vol. 36, pp. 554-558, 2011.

[8] M. Takino, "Use of liquid chromatography/mass spectrometry for the pesticide residue analysis: pesticide residue analysis in food by LC-MS/MS and LC-Q/TOFMS," J. Pestic. Sci., vol. 37, pp. 279-302, 2012.

[9] D. Liang, W. Liu, R. Raza, Y. Bai, and H. Liu, "Applications of solid-phase micro- extraction with mass spectrometryin pesticide analysis," J. Sep. Sci., vol. 42, pp. 330-341, January 2019.

[10] J. Wong, J. Wang, W. Chow, R. Carlson, Z. Jia, K. Zhang, et. al, "Perspectives on liquid chromatography-high-resolution mass spectrometry for pesticide screening in foods," J. Agric. Food Chem., vol. 66, pp. 9573-9581, September 2018.

[11] I. Lee, J. Lee, "Effects of thermal aging on cellulose pressboard using terahertz time-domain spectroscopy," Curr. Appl. Phys., vol. 19, pp. 1145-1149, November 2019.

[12] D. Suzuki, Y. Kawano, "Terahertz imaging and spectroscopy as a tool for non-destructive and non-contact quality inspections of medical

drugs and polymer films," Bunseki Kagaku, vol. 66, pp. 893-899, December 2017.

[13] H. Yudistira, L. Yehezkiel, and K. Kananda, "High absorbance performance of symmetrical split ring resonator (SRR) terahertz metamaterial based on paper as spacer," Mater. Res. Express, vol 6, pp. 02584, February 2019.

[14] A. Rahman, A. K. Rahman, "Terahertz technology enables systems for molecular characterization," Laser Focus World, vol. 1, pp. 113–117, January 2012.

[15] E. Grercht, L. You, "Terahertz imaging and spectroscopy based on hot electron bolometer (HEB) heterodyne detection - art. no. 689308," vol. 6893, pp. 89308-89308, January 2008 [Conference on Terahertz Technology and Applications, 2008].

[16] J. Ornik, D. Knoth, M. Koch, and C. Keck, "Terahertz-spectroscopy for non-destructive determination of crystallinity of L-tartaric acid in smartFilms (R) and tablets made from paper," Int. J. Pharm., vol. 581, pp. 119253, May 2020.

[17] M. Wolfgang, A. Peter, P. Wahl, D. Markl, J. Zeitler, and J. Khinast, "At-line validation of optical coherence tomography as in-line/at-line coating thickness measurement method," Int. J. Pharm., vol. 572, pp. 118766, December 2019.

[18] M. P. Wagh, Y. H. Sonawane, and O. U. Joshi, "Terahertz technology: a boon to tablet analysis," Indian J. Pharm. Sci., vol. 71, pp. 235-241, May 2009.

[19] J. Lin, T. Yen, T. Huang "Design of annulus-based dielectric metamaterial cloak with properties of illusion optics," J. Opt., vol. 22, pp. 085101, Augest 2020.

[20] K. Deng, Y. Ding, Z. He, H. Zhao, J. Shi, and Z. Liu, "Theoretical study of subwavelength imaging by acoustic metamaterial slabs," J. Appl. Phys., vol. 105, pp. 124909, June 2009.

[21] N. Kundtz, D. R. Smith, "Extreme-angle broadband metamaterial lens," Nat. Mater., vol. 9, pp. 129-132, February 2010.

[22] W. Liu, F. Fan, S. Chang, J. Hou, M. Chen, X. Wang, and J. Bai "Nanoparticles doped film sensing based on terahertz metamaterials," Opt. Commun., vol. 405, pp. 17-21, December 2017.

[23] Y. Wang, Z. Cui, D. Zhu, L. Yue, "Composite metamaterials for thz perfect absorption," Phys. Status Solidi A, vol. 216, pp. 1800940, March 2019.

[24] Y. Wang, D. Zhu, Z. Cui, L.Hou, L. Lin, F. Qu, et. al, "All-dielectric terahertz plasmonic metamaterial absorbers and high-sensitivity sensing," ACS Omega, vol. 4, pp. 18645-18652, November 2019.

Super-Resolution Imaging Using Very Deep Convolutional Network in Terahertz NDT Field

Qiang Wang
Air Force Engineering University
Xi'an, China
caption_wang@21cn.com

Hongbin Zhou
Air Force Engineering University
Xi'an, China
hbzhou990@126.com

Yi Wang
Air Force Engineering University
Xi'an, China
61846548@qq.com

Ruicong Xia
Air Force Engineering University
Xi'an, China
rayc0-xia@foxmail.com

Qiuhan Liu
Air Force Engineering University
Xi'an, China
1178556035@qq.com

Boyan Zhao
Air Force Engineering University
Xi'an, China
394762422@qq.com

Yue Wang *
Air Force Engineering University
Xi'an, China
wangyue2017@xaut.edu.cn

Abstract—In this paper, we propose a Terahertz image Super-Resolution (SR) method based on very deep Convolutional Networks (VDCN) in non-destructive testing (NDT) application. Glass fiber reinforced plastic (GFRP) with prefabrication defects is used as test specimen. Through fast scanning imaging by Terahertz domain spectrometer, we obtained the single original image. Then the HR image can be generated by trained VDCN. Experimental results show that the image restored by our method has higher peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) than traditional Bicubic.

Keywords—Terahertz image, Super-Resolution imaging, Convolution network

I. Introduction

Terahertz (THz) radiation refers to the electromagnetic wave whose frequency ranging from 0.1~10THz. As the most promising Terahertz technology, active Terahertz imaging [1-3] has been applied in Non-destructive testing (NDT) [4-5,14], security scanning [6-7] and other fields. Due to the low photon energy, Terahertz image has low signal-to-noise ratio (SNR), serious blur, low resolution, which usually does not meet the requirement of human's vision [8]. Therefore, effective methods are demanded to improve the resolution and SNR of Terahertz image.

Super-Resolution algorithms (SRA) addressed the problem of using low-resolution (LR) image generating high-resolution (HR) image commonly requires more image details [9]. SRA has been studied for many years in the computer vision community. Preceding methods include interpolation (e.g. Bicubic interpolation), Lanczos resampling [10], and more superior ones utilizing statistical image priors [11-12] or internal patch recurrence [13].

A Super-Resolution Terahertz imaging method applying on NDT using very deep Convolutional Networks is proposed in this paper to generate HR Terahertz image with high SNR. Glass fiber reinforced plastic (GFRP) with prefabrication defects was used as test specimen. Through fast scanning imaging by Terahertz domain spectrometer, we obtained the single original image. Then the HR image can be generated by trained very deep convolutional networks. Experimental results show that the image restored by our method has higher peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) than traditional Bicubic interpolation.

II. Related Work

Kim et al. presented a high-precision single image super-resolution (SISR) method was proposed, using a very deep network called VDSR, and then they found that increasing the network depth can significantly improve the accuracy [9].

Liu et al. presented an improved convolutional neural network (CNN) structure based on channel combination, which can not only train quickly but also improve the accuracy. In addition, they proposed to extract feature maps and an efficient multi-channel convolution layer in LR space, which learned a series of upper scale filter arrays specially trained for each feature map, so as to enhance the final HR feature mapping to the HR output [15].

Liu et al. proposed a Terahertz NR lens by using a four-wave mixing process in graphene. Theoretical analysis and numerical simulations were performed to demonstrate the capability of such imaging [16].

Xie et al. proposed an adaptive super-resolution reconstruction method for terahertz image based on Markov random field. This method not only has high super-resolution, but also can better preserve the edge information of the image and reduce the noise of the restored image [8].

III. Deep Learning Architecture

Here CNN was used for Terahertz image SR reconstruction. Fig. 1 shows the structure of VDCN adopt in this paper which has been proved to be a useful and quick networks for SISR [9]. Based on this network, we added a Terahertz image process layer for image size adjustment and channel change. Layers were designed the same structure: 64 filters with input size $3 \times 3 \times 64$ except the first and the last layer, where a filter operates on a 3×3 spatial region across 64 channels. Image reconstruction was designed in the last layer, which contains single filter with $3 \times 3 \times 64$ size.
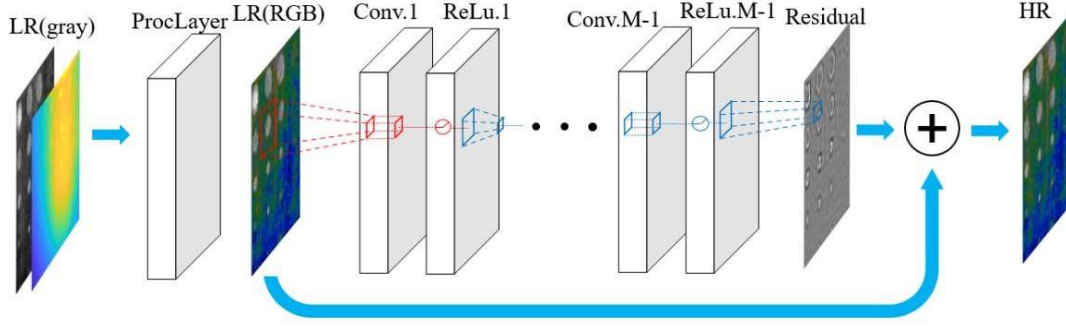
Fig. 1. Structure of VSDN used in our Terahertz SISR experiments

(a) peak-to-peak   (b) max amplitude   (c) 0.16 THz
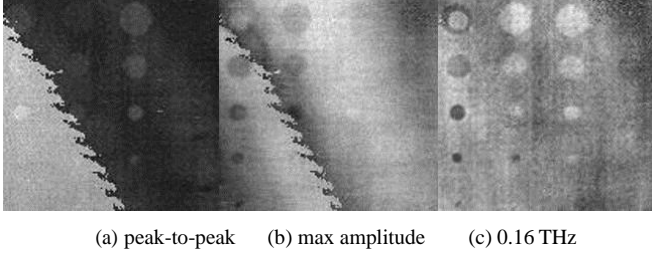
Fig. 2. Results of different Terahertz imaging methods

The network uses low resolution interpolation image as input to predict image details. Super-Resolution methods are often use modelling image details and we found that CNN-based method performs well in this field. However, using VDCN to generate HR image brought a problem that the size of the feature map decrease every-time when convolution operates.

The random initialization method is used to determine the weight, and the back-propagation calculation is used to update

the network weight, so that the loss value is reduced to stable.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. Dataset

In order to finish experiment in the Terahertz SR imaging using the VDCN, more composite structures (e.g. laminate, A sandwich and C sandwich structure) with embedded defects were imaged. Besides, as for the same sample, we used various Terahertz imaging methods, including max amplitude, flight-time, peak-to-peak so on, for generating discriminated images.

Fig. 2 shows Terahertz images obtained by three methods (peak-to-peak, max amplitude and 0.16 THz frequency domain imaging method) of GFRP with embedded defects. The detail of imaging GFRP sample is shown in Fig. 3, which is consist of four laminates with 20 mm as defect max-diameter and 3 mm as defect min-diameter. And all defects are distributed at four depths as to 0.25 mm, 0.5 mm, 0.75 mm, 1 mm.

B. Data Preprocessing

The VDCN seeks to learn implicit redundancy that present in natural data to recover missing HR information from a single LR instance [15]. However, Terahertz imaging usually
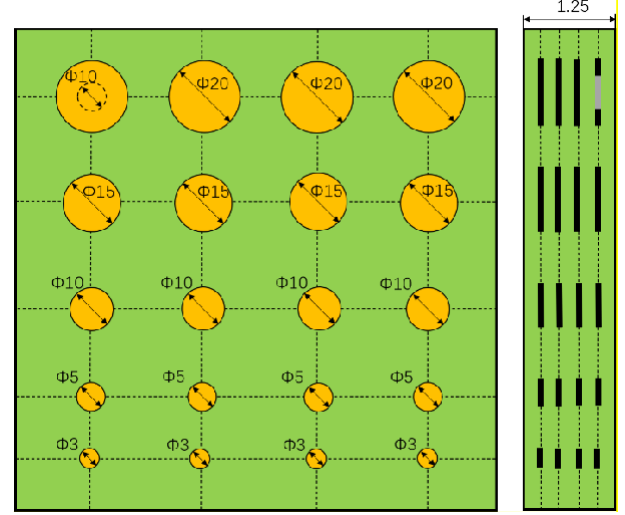
Fig. 3. Sample design drawing

TABLE I.   DATA DISTRIBUTION FOR TRAINING AND TESTING

	Training Set	Testing Set
Peak-to-Peak	121	20
Max Amplitude	130	20
Flight Time	145	20
Frequency	144	30

generates gray image. Thus, we proposed a method for transforming gray image to RGB image. First, the gray image is expanded to three channels by copying channels where an another RGB image should be provided for reference. Thermography is used for color match in this paper. Then the two images were converted into YCbCr space. The result graph is generated pixel by pixel through a double loop.

Traditional data extension methods usually focus only on the detected object itself, with little consideration for background information. In fact, the background information needs to be considered in the training calculation, because the characteristics of the object need to be extracted from the complex background. In this paper, a data enhancement method based on background information is proposed to improve the generalization performance of VDCN, and the process is showed in Figure 4. The number of data set is shown in Tab. 1.

(a) Scale Factor ×2   (b) Scale Factor ×3   (c) Scale Factor ×4
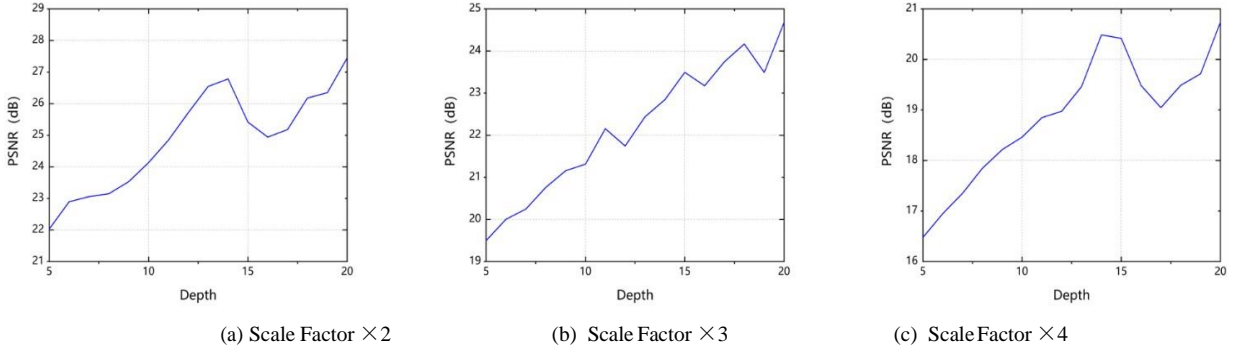
Fig. 4. Depth vs. Performance
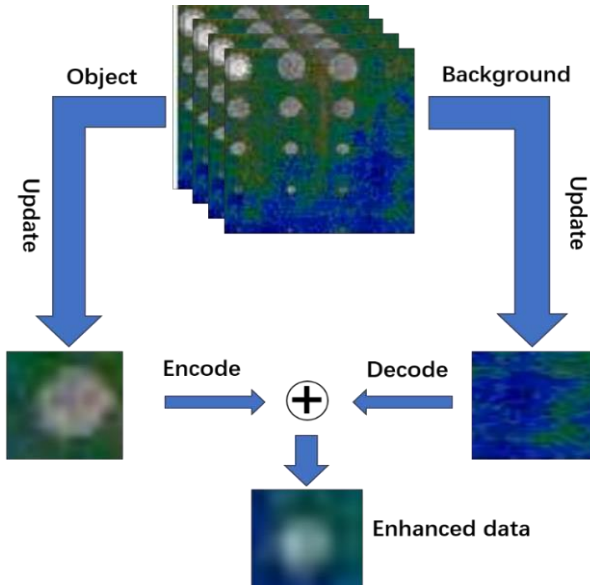
Fig. 5. The data enhancement process based on the fusion of object and background

C. Results and Discussion

A small batch gradient descent method based on back propagation is used to optimize the regression target for network training. To avoid the local optimal solution, we set the learning rate as 0.0001. In order to obtain enough defect image features, batch and subdivisions were set to 128 and 32.

In order to verify the efficacy of VDCN performance on SR, we trained and tested networks with depths between 5 and 20. Fig. 5 shows the results, performance was promoted rapidly as depth raised.

The residual network in this work was used to learn Terahertz images and compared with non-residual network experimentally. Set depth 10 (weight layer) and scale factor 2. The same learning rate scheduling mechanism is used. To our surprise, the residual network shows superior performance and faster speed in convergence as shown in Fig. 6.

Lastly, we compared our method with Bicubic SR and original image. The result is shown in Fig. 7. Defects of depth 1mm could not been seen in all results even by SR images. However, defects of diameter 3mm could be identified through our method which was invisible in Bicubic and original images. Thus it can be seen that SR method could realize smaller defect detection, but it was not helpful to more
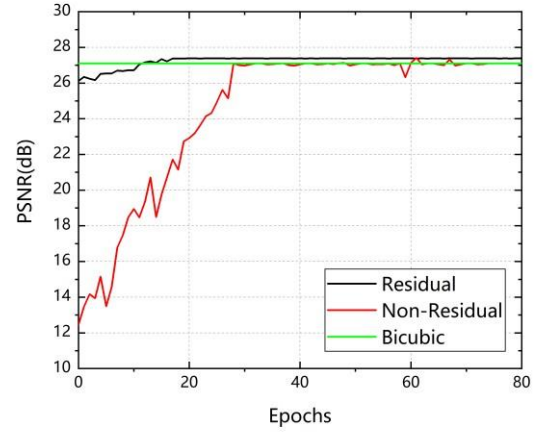
Fig. 6. PSNR variation between three networks

TABLE II.   PSNR/SSIM VALUE OF SCALE FACTOR 2, 3, 4. RED COLOR INDICATES THE BETTER PERFORMANCE

	Scale Factor	VDCN	Bicubic
PNSR	×2	27.3836	27.1012
	×3	24.8164	24.0682
	×4	20.7201	20.5992
SSIM	×2	0.9519	0.9506
	×3	0.9365	0.9224
	×4	0.8278	0.8282

deeper ones. In Tab. 2, we summarized the quantitative evaluation of the two SR methods. When the scale factor is 2, 3, 4, VDCN is better than Bicubic. Nevertheless, the performance of VDCN is narrowly better than Bicubic. The main reason is that Terahertz image and natural image differentiate significantly. Besides, the network parameters were not adjusted to excellent status.

V. CONCLUSIONS

In this work, Terahertz SR imaging using VDCN for improving the capability of Terahertz NDT was realized. For VDCN demanding RGB image, we proposed an image process method to transform original Terahertz images which are in gray color to RGB ones. Besides, taking background information into consideration, the enhanced image was fused
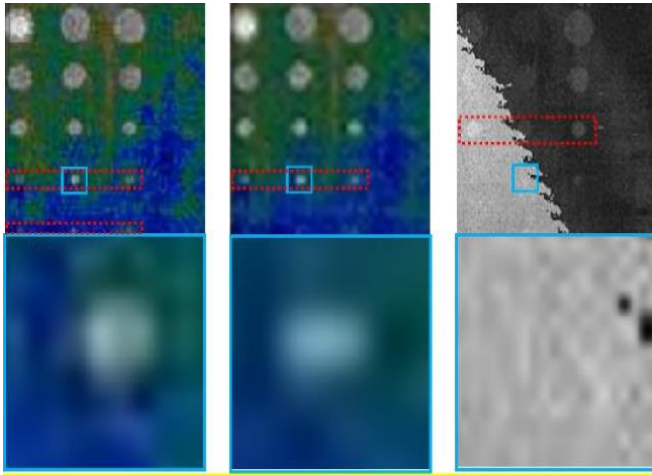
VDCN(24.8164, 0.9365) Bicubic(24.0682, 0.9224) Origin(13.1494, 1.0000)

Fig. 7. Comparison of SR results of Terahertz image with scale factor ×3 with original image

by objects and background extracted from different Terahertz images respectively. The experiment showed that deeper networks attained better efficacy. In addition, residual networks exhibited superior performance and faster speed than non-residual networks at convergence. Lastly, compared with Bicubic and original images, the proposed method outstood slightly in SR imaging for Terahertz image which performed well in smaller defect. Internal cracks are more common in composite than delamination. We hope to detect small cracks by SR imaging. Whereas, numerous Terahertz image data and benign methods for network optimization are required to improve the SR efficacy in Terahertz NDT field which could be the research focus in the future.

REFERENCES

[1] Hu B. B., Nuss M. C. IMAGING WITH TERAHERTZ WAVES[J]. Optics Letters, 1995, 20(16): 1716-1718.

[2] Mittleman D. M., Jacobsen R. H., Nuss M. C. T-ray imaging[J]. IEEE Journal of Selected Topics in Quantum Electronics, 1996, 2(3): 679-692.

[3] Zhang X. C. Terahertz wave imaging: horizons and hurdles[J]. Physics in Medicine and Biology, 2002, 47(21): 3667-3677.

[4] Stoik C. D., Bohn M. J., Blackshire J. L. Nondestructive evaluation of aircraft composites using transmissive Terahertz time domain spectroscopy[J]. Optics Express, 2008, 16(21): 17039-17051.

[5] Cristofani E., Friederich F., Wohnsiedler S., et al. Nondestructive testing potential evaluation of a Terahertz frequency-modulated continuous-wave imager for composite materials inspection[J]. Optical Engineering, 2014, 53(3): 031211.

[6] Kemp M. C., Taday P. F., Cole B. E., et al. Security applications of Terahertz technology[C]. //Proceedings of the Terahertz for Military and Security Applications, 2003: 44-52.

[7] Zimdars D., White J., Williamson S., et al. High-speed time domain Terahertz security imaging[C]. //Proceedings of the Terahertz for Military and Security Applications III, 2005: 131-136.

[8] Xie Y Y , Hu C H , Shi B , et al. An Adaptive Super-Resolution Reconstruction for Terahertz Image Based on MRF Model[J]. Applied Mechanics and Materials, 2013, 373-375:541-546.

[9] Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee. Accurate Image Super-Resolution Using Very Deep Convolutional Networks[C]. // 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1646-1654.

[10] C. E. Duchon. Lanczos filtering in one and two dimensions.Journal of Applied Meteorology, 18(8):1016–1022, 1979.

[11] J. Sun, Z. Xu, and H.-Y . Shum. Image Super-Resolution using gradient profile prior. In CVPR, 2008.

[12] K. I. Kim and Y . Kwon. Single-image Super-Resolution using sparse regression and natural image prior. TPAMI, 2010.

[13] D. Glasner, S. Bagon, and M. Irani. Super-Resolution from a single image. In ICCV, 2009.

[14] Wang Q , Li X , Chang T , et al. Nondestructive imaging of hidden defects in aircraft sandwich composites using Terahertz time-domain spectroscopy[J]. Infrared Physics & Technology, 2019.

[15] Liu C, Li Y, Luo J, et al. A novel convolutional neural network architecture for image Super-Resolution based on channels combination[C]// International Conference on Information Fusion. IEEE, 2017.

[16] Jiang-Yu L, Tie-Jun H, Pu-Kun L. Terahertz Super-Resolution imaging using four-wave mixing in graphene[J]. Optics Letters, 2018, 43(9):2102.

Generative Model-Driven Sampling Strategy for High Efficient Measurement of Complex Surfaces on Coordinate Measuring Machines

Jieji Ren[*]
School of Mechanical Engineering
Shanghai Jiao Tong University
Shanghai, China
jiejiren@sjtu.edu.cn

Mingjun Ren
School of Mechanical Engineering
Shanghai Jiao Tong University
Shanghai, China
renmj@sjtu.edu.cn

Lijian Sun
AI Research Institute
Zhejiang Lab
Hanzhou, China
sunlj@zhejianglab.com

Abstract—**Coordinate measuring machine is widely used in precision measurement of industrial parts. However, the nature of the point-by-point probing characteristic limits its efficiency in the measurement of complex parts which normally requires dense sampling pints for fully evaluating the machining errors with high fidelity. To address this problem, this paper proposes a generative model-driven sampling strategy to reduce the number of the sampling points while maintaining the measurement accuracy. Specifically, the surface error reconstruction under sparse sampling is transformed as an image super-resolution task, which adopts a generative model to estimate accurate dense results from under-sampled data. A multi-scale neural network architecture is designed to achieve reconstruction, and the Fractional Brownian Motion is applied to synthesis large-scale simulated error datasets for model training. The generalized neural model could use sparse measurements to reconstruct global machining error, which dramatically reduces the sampling time and increases measurement efficiency. Both computer simulation and actual measurement are carried out to verify the effectiveness of the proposed method.**

Keywords—Machining Error, Measurement, Deep Learning

I. INTRODUCTION

Machining error evaluation is significant in industry, especially in advanced manufacturing[1]. There are generally two main measurement approaches for complex surface reconstruction and error evaluation[2]. Contactless approaches, such as optical probe, structure lights and high-accurate vision system[3], could achieve high efficiency, but they have strict requirements for measuring surfaces and are difficult to measure shiny surfaces. Furthermore, these methods are affected by environment and may introduce large noise and measurement error, which limits contactless measurement applications. On the other hand, contact approaches, such as CMM and surface profile meter could achieve high accuracy by using probe contact with surface and recording coordinates[3]. The measurement accuracy is remarkable, especially in precision measurement, but point-by-point sampling manner retard the whole process. Because of the physical principle of CMM and the accuracy requirements of the control system, the speed of the measuring probe was limited, and the sampling process is less efficient. The typical time for precision testing with a coordinate measuring machine is usually on the order of 10^0-10^2 minutes [4]. This contradictory is more severe in large-scale complex surface measurements.

To accelerate the measurement process, various sampling and reconstructing strategies are proposed, which reduces measure time by sampling sparsely, and reconstructs completed results by utilizing prior or statistics of surfaces[5].

Reducing the number of sampling points is an effective way to increase the measurement efficiency. Nevertheless, reconstructing accurate surfaces with limited samples is one of the core problems in the precision measurement field. Since the machining error scale is smaller than the surface scale, it poses more severe challenges for precise machining error reconstruction.

Researchers have proposed a variety of reconstruction methods to reconstruct accurate machining error with limited samples[6]. Machining error reconstruction is a process of building surface mathematical model under sparse measurement, which includes explicit and implicit methods. For explicit methods, parameterized B-spline[7],[8] is commonly used for surface fitting. NURBS[9] based degrees of freedom, control points, knot vector and weight factors to fit surface, which provides a general and accurate representation for free surface and is invariant under affine or perspective transformations. Triangulation methods[10], such as Delaunay triangular (DT), utilized triangular patches to fit surface. Different curvatures of surface could be handle by various size triangle patches[11]. DT could reconstruct surface with uniqueness, proximity, optimality, and convex polygons. These methods are general used in complex surface reconstruction. However, the reconstruction precision degrades dramatically as measurement points decrease.

For implicit methods, it uses implicit functions, consists with various basic functions, to model the surface[12]. Least square(LS) is a generally used surface fitting method. Besides, radical basis function(RBF)[13] and kernel function regression methods are also utilized to reconstruction. Recently, Bayesian methods, especially Gaussian Process Regression(GPR), achieved success in complex surface reconstruction with CMM[14] and metrology with non-raster scanning probe microscopy, with sparse measurement[15]. It bases Bayesian prior and utilizes correlation of sparse measurements to reconstruct surface with high precision and flexibility. Although reconstruction accuracy is impressive, covariance and kernel functions introduce matrix inversion, which has high computational complexity as $O(n^3)$, even longer than sampling[16]. The intense computation makes it nearly impossible to apply to large-scale measurement, in the practical production process. Besides, hyper-parameters need verbose manual tuning, which is time-consuming and highly depends on experience. These reasons limit its application to efficient measurement.

In recent years, as the big data resource and computing power increases, deep learning methods achieve great success on various computer vision task[17]. Among them, image super-resolution(SR) task provide a feasible solution for fast and precise machining error reconstruction with sparse measurement[18], which uses neural network to

---

*Jieji Ren is the corresponding author. (e-mail: jiejiren@sjtu.edu.cn).

reconstruct the high-resolution(HR) images from the low-resolution(LR) input.

SRCNN[18] first address this problem by neural network. With 3 layers structure, SRCNN extract features from LR image, and mapping these features into HR feature space to reconstruct HR image. Then, FSRCNN[19] adopts deconvolution layer to up-sampling and reconstruct HR image directly from LR inputs, replaced preprocess in SRCNN. ESPCN[20] uses extra channels to generate more information by increase features map. Sub-pixel convolution technique decreases kernel size and computes complexity. VDSR[21] explores very deep model in SR, applies multi-scales and residuals to improve reconstruction performance. DRCN[22] uses same kernel to reduce parameters and computation. SRResNet[23] uses multi-layer residual to increase neural network depth. DRRN[24] uses recurrent block to reuse parameters, and EDSR[25] uses more channels to increase train process and multi-scale adaption. Also, SRDenseNet[26] uses dense connected to increase super-resolution reconstruction results.

The key of image super resolution is learning a model to map the sparse measurements into dense measurement with accuracy[27]. This process is quite similar to reconstruct ME under down-sampled measurement. Inspired by SISR, we consider ME reconstruction with sparse measurements as an image super-resolution task, which could leverage advanced CNNs to improve error evaluation performance and utilized GPU to accelerate this process.

To improve machining error evaluation efficiency, this paper proposes a **high-efficiency and generative model-driven measurement strategy**, which could efficiently reconstruct precise machining error. Firstly, we use fractal Brownian motion (fBm)[28] to effectively model and describe ME[29]. Then, we build a novel multi-scale neural networks to reconstruct completed dense ME with sparse measurement inputs. To train networks, we introduce fBm to synthesis dense ME data and apply down-sampling to generate sparsely sampled measurements, which provides a paired data set for the neural model to learn sparse-to-dense reconstruction. In practice, down-sampled measurements of a workpiece are compared with designed truth, to get sparse error data. Sparse machining errors are arranged into array and sent into model to reconstruct accurate dense ME. **The proposed strategy dramatically reduces sampling time while preserves high measurement accuracy.** Simulation results and practical measurement verified the effectiveness and generalization of the proposed method.

II. PROPOSED METHOD

This paper proposes generative model-driven strategy to achieve fast machining error reconstruction of the complex surface under sparse measurement from CMM. To accelerate surface machining error evaluate process, we first used a sparse sampling strategy to reduce the time consumption of measure. Then, sparse measurements subtract with design truth of the surface to generate sparse machining error. These sparse errors are sent into reconstruction neural networks to generate dense machining error. This work adopts auto-encoder based neural network as reconstruction architecture, as a generative model, auto-encoder could extract the surface features from training data and learn the mapping between sparse measurement and dense results. This strategy provides a new perspective for surface error evaluation and metrology.
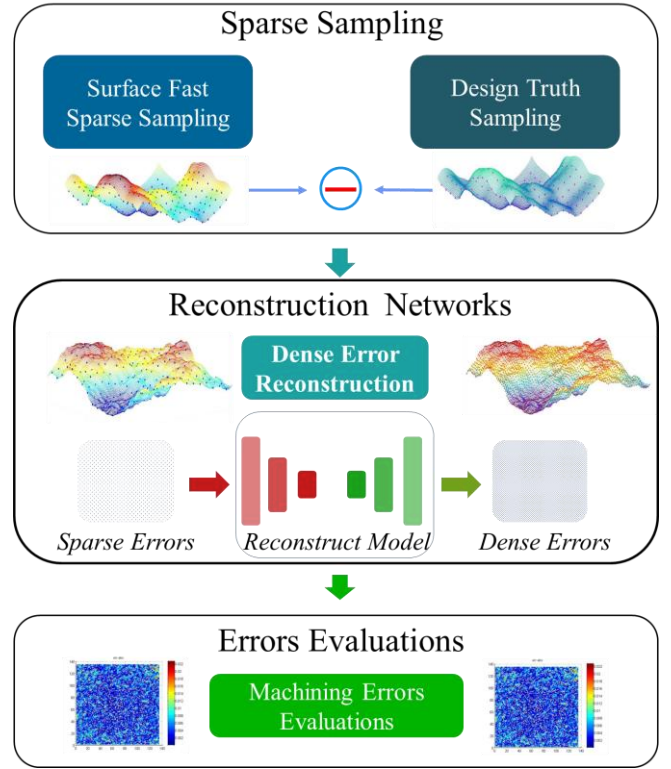
Fig.1. Framework of proposed methods, which utilitzed generative model to reconstruct dense measurements from sparse samples.

This model could learn the mapping relationship between sparse sampling and dense reconstruction results. To train the model, we synthesis simulation data by fBm, this could approximately describe surface machining error. With this model, we could take down-sampling measurement as inputs, and inferred the high accurate surface machining error reconstruction result in real-time. Fig.1. shows the entire pipeline.

A. Mahining Error and Fractal Brownian Motion

Machining error are inevitable in machining process[30]. These errors result from random motion between cutting tool and machining workpiece, machine tool intrinsic error, errors from tools fixtures, workpiece thermal and mechanical deformation and internal stress[29]. These reasons would lead actual machining surface deviate from designs.

To model machining error, we suppose the relative positioning error of cutting tool and machining workpiece is random noise with Gaussian distribution, the machining error can be considered as a signal generated by random walks, which increments satisfied Gaussian distribution. Hence, in this study, fractional Brownian motion (fBm) is used to generate the fractal surface to simulate the machining error of the machined surface. Because fBm could describes complex surface situation, many researches applied brownian surface(BS)[4] simulated terrain of earth and surface of machined workpieces. fBm is the basis of brownian surface, which is a continuous-time Gaussian process $B_H(t)$ on $[o; T]$ with following covariance function:

$$[B_H(t)B_H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}) \qquad (1)$$

Where H is Hurst parameter, which describes the roughness of motion process, the higher H could leads smoother results.

For three-dimension, we could apply fBm as elevation function with x, y coordinates, the result is BS. BS is natural looking surface, we could generate different roughness surface by adjusting H parameter.

$$E[B_H(x)B_H(y)] = \frac{1}{2}(|x|^{2H} + |y|^{2H} - |x-y|^{2H}) \qquad (2)$$

Here, $x, y \in [0,T]$. We simulated surface with different Hurst index, as figure xx shows.
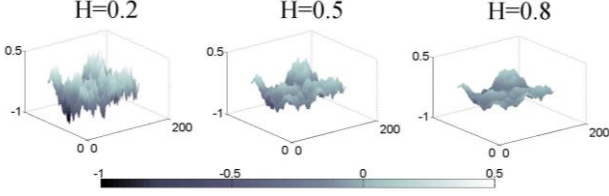
Fig.2. Generated Brownian surface with different Hurst parameters

When the H=0.8, the surface exhibits strong low-frequency characters and middle scale error, which match the machining error process. In this work, we set Hurst index as 0.8 for surface generation, as simulated training dataset. We use various initial parameters to generate Brownian surface, which are enough to cover possible machining errors, for model learning.

B. Reconstruction from Sparse Measurement

Reconstruct dense accurate surface from under-sampling measurements could be considered as an nonlinear mapping problem. We need to transfer measurements from under-sampling space into dense-sampling space. This is a high nonlinear and complex task, so we take neural network to modeling this process. For example, if the original sparse measurement with $rN * rN$ points, the under-sampling measurement with $N * N$ points, we need reconstruct $rN * rN$ results based $N * N$ measurements, where r is down-sampled rate.

Because we want to model the sparse-dense mapping process in same scale(as measurement physical process), we arrange sparse data into array, which size same as dense measurement array. For example, if we want to generate $rN*rN$ dense measurement from sparse $N*N$ grid sampling, we firstly put $N*N$ points into $rN*rN$ array with corresponds x,y coordinates, the un-sampled locations will set as 0, as placeholder. Then, we will send the rearranged sparse samples into model to complete missing values and generate
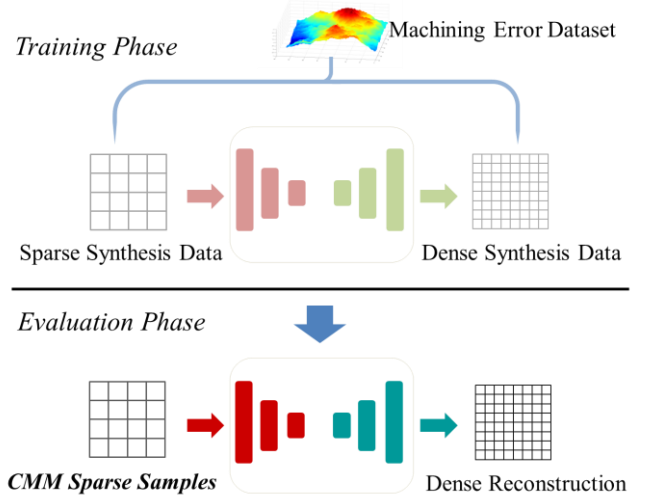
Fig.3. Pipeline of reconstruction process. Synthesis dataset is used to train model. Then, sparse measure errors from CMM could be used to reconstruct dense machining errors.

$rN*rN$ dense results. It is worth mentioning that we use regular raster down-sampled strategy to reduce measurement points, which is suit for neural network to process.

C. Auto-encoder based Reconstruction Archetecture

In order to reconstruct dense results from sparse samples, we base auto encoder-decoder structure and convolutional neural network(CNNs) build our neural network. CNNs could extract multi-level features from input, and have enough representation ability to describe the accuracy mapping process. We first use encoder to extract data features from inputs, which not only contains global trend, but also describes local details. Model could learn the properties of machining error and generate implicit mapping relations from sparse inputs and dense results. Based on these features, decoder could generate dense machining error on the un-sampled location. To facilitate the flow of sparse features, skip connections are used to link same level of encoder and decoder. Furthermore, to improve the stability of the network and accelerate training convergence, we utilized multi-scale branches to reconstruct 1/2 and 1/4 scale results.

Model needs to learn the mapping relations between sparse inputs and dense measurement results. The dense measure reconstruction process could express as following：

$$M_{dense} = \mathcal{F}\left(M_{sparse}; \theta\right), \qquad m(x, y, z) \in \mathcal{R}^3 \qquad (3)$$
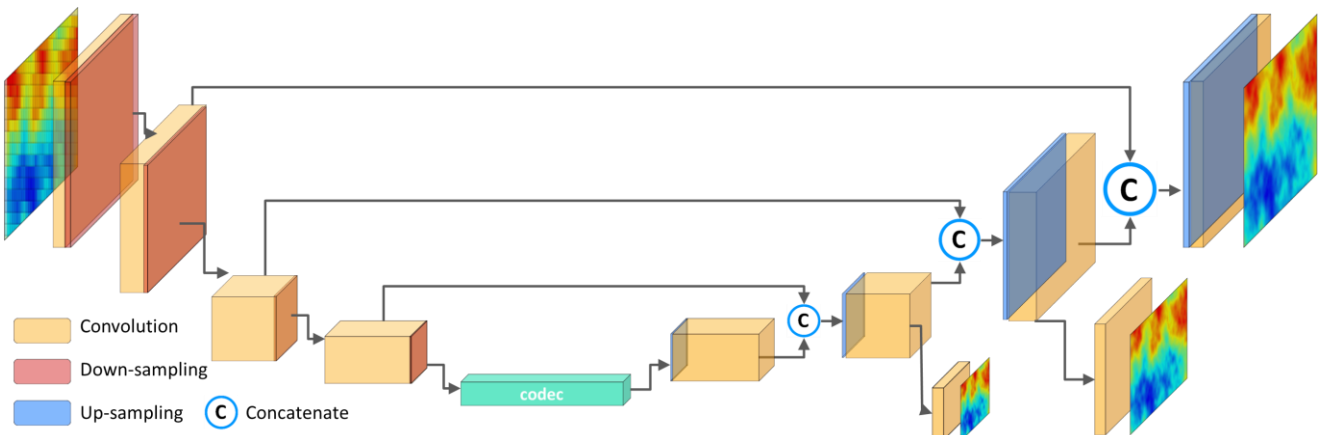
Fig.4. Architecture of reconstruction neural networks.

Model should guarantee point-wise accuracy and global performance. In machining error reconstruction, RMSE and error peak-value are important to measurement evaluation, express as bellow:

$$MSE = \frac{1}{N^2} \sum_{i,j=1}^{N} \left\| Y_{ij} - \overline{Y}_{ij} \right\|^2$$

$$RMSE = \sqrt{\frac{1}{N^2} \sum_{i,j=1}^{N} \left\| Y_{ij} - \overline{Y}_{ij} \right\|^2} \qquad (4)$$

$$PV = \max \left| Y - \overline{Y} \right|$$

Where Y and $\overline{Y}$ are ground truth and predict results. The loss function of specific scale reconstruction could write as following:

$$L = \frac{1}{N^2} \sum_{i=1}^{N} \left\| Y_i - \overline{Y}_i \right\|^2 + \alpha * \max \left| Y - \overline{Y} \right|$$
$$L = L_{mse} + \alpha * pv \qquad (5)$$

L_{mse} drive model to improve point-wise accuracy, supervised by dense accurate measurement. To avoid noise or local extreme points affect global performance, we also introduced p - v (peak-valley) value as part of loss function. Parameters α controls p-v contribution to training.

Loss of 1/2,1/4 scale are same as previous, while the ground truth are down-sampled from original scale ground truth. Whole networks optimization could express as following:

$$\arg \min_{\theta} \left\{ \sum_{scale} L_{mse} + \alpha * pv \right\} \qquad (6)$$

To strengthen the original scale reconstruction, loss of [1,1/2,1/4] scale could multiply weights as [1,0.5,025].

D. Data Synthesis and Training Process

To train this model, we syntheses large-scale dataset with previous method. This dataset contain 60,000 surface error samples, every sample contains paired dense accurate surface error and under-sampled measurements.

For the dense measurements, we generate $rN*rN$ data points array, and used regular down-sampled method to obtain $N*N$ sparse measurements, with sparse factor $1/r$. With this strategy, numbers of sample could reduce r^2 times, which could increase efficiency by r^2 times on measurement phase. While the reconstruction phrase utilized neural network to process, which could be accelerated by GPUs.

This paper chooses $N=[64,128,256]$ to cover different resolution and set sparse factor $r=[4,8,16]$ to increase generality of algorithm. The model takes down-sampled sparse data as inputs, and reconstructs dense measurement array as goal. Training process will make model obtain the ability that reconstruct dense error from sparse measurement.
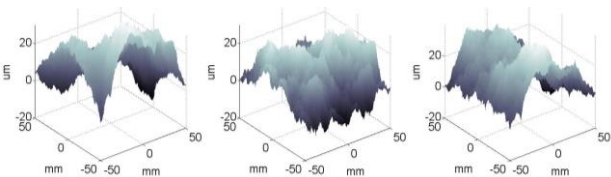
Fig. 5. 5Synthesis surface machining error samples with fBm.
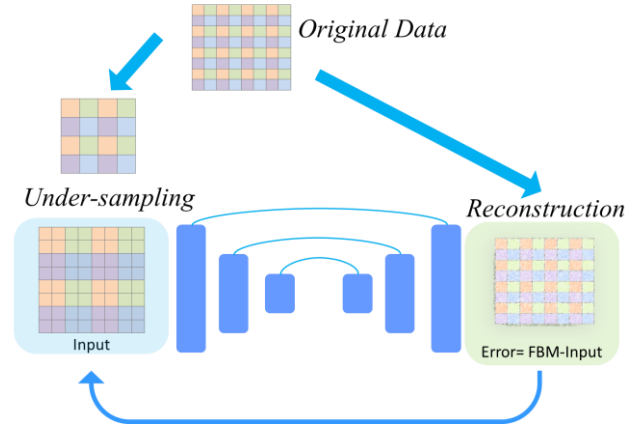
Fig.6. Model training process, which take under-sampled measurements as inputs, and reconstruct dense machining error as goal.

Furthermore, to evaluate performance of this model quantitatively, the horizon range of synthesis machining error is set to [-50mm,50mm], the height range are normalized into100um.(normalization on training)

This work uses Nvidia V100 GPU to train, the initial learning rate is 10^{-4}, weight-decay is 0.999, the p-v loss weights α is 0.03, batch size is 64, training epoch is 100. The model is built by Tensorflow framework with python.

III. EXPERIMENT

Model is training on synthesis dataset and test on machined work-piece. Firstly, synthesis dataset set $N=128$ grid data as objective dense reconstruction, while sample rate are set as $r=[2,4,8,16]$, which could reduce the number of measurement points into [1/4,1/16,1/64,1/256]. Model take sparse measurements as inputs and learning to reconstruct dense results. Training process is shown as Figure.7.
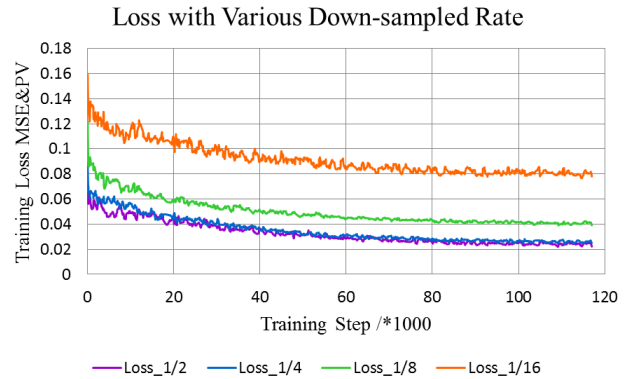
Fig.7. Model training process, surface reconstruction error, such as MSE and p-v values, decreases with training epochs. Training data are normalized.

To fitting different down-sampling scale, model utilizes datasets under various sample rate $r=[2,4,8,16]$ for training and testing, which explore the robustness of this approach under various measurement sparsity.

Fig.8 shows reconstruction accuracy changes with real under sampled rate [4,16,64,256]. This method is able to restore dense surface machining error reliably. The accuracy just slightly decline with the sampling rate decrease, which further verified effectiveness and generalization of this methods, without any changes or tuning. Apart from regular raster sampling, we could adopt various sampling rates and strategies in training phase, such as random and adaptive sampling, to increase generality and adaptability of model.
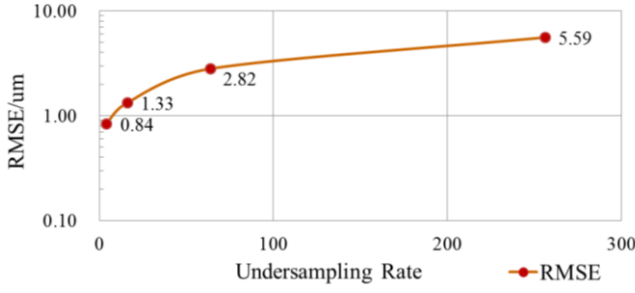
Fig.8. Reconstruction accuracy change with measurement sampling rate. Proposed mehtod shows robustness, that accuracy slightly decline with dramatically drop of undersampling rate from *4* to *256*.

To show the reconstruction performance of this strategy and compare with other methods, results under $r=16$ are drew in Fig.9. The reconstruct surface error is extremely close to original results. It also could reconstruct details in rapidly changing regions, such as peaks and valleys.
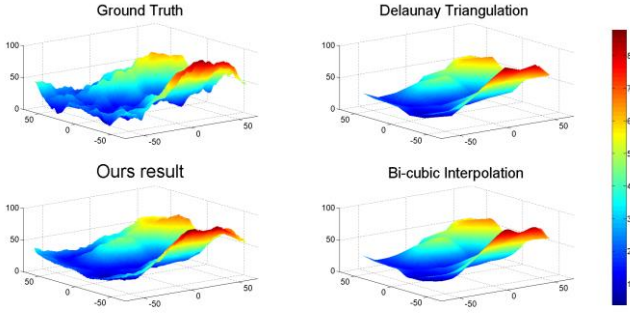
Fig.9. Reconstruction results. Our method could restore in detail, better than interpolated methods.

We also compared our results with interpolated method, Delaunay Triangulation and GPR among 1,000 samples. The reconstruction errors metrics and times are recorded. Tab.1 shows proposed approach is fast and accurate.

TABLE I.   RECONSTRUCTION RESULTS WITH DIFFERENT METHODS

Method/Metrics	Ours	Bi-Cu	Delau	GPR
P-V(um)	23.95	126.1	123.8	150.3
RMSE(um)	5.447	24.05	23.54	46.40
Time(ms)	3.4	4.1	9.5	1163

Compare with Bi-cubic interpolation, our method could achieve more accurate reconstruct results, the p-v and RMSE are smaller than Bi-cubic methods. Compare with Gaussian Process Regression method, our method is much faster without hyper-parameters tuning. Furthermore, neural network could accelerate by GPUs while GPR is not easy to increase efficiency by parallelization. Proposed approach is surface-agnostic methods, which could applicate to various machining error reconstruct with corresponding training data. It is worth to note that when the sampling rate is extremely sparse, 1/256 in this experiment, interpolation based methods decline dramatically. Without fine hyper-parameters tuning, GPR methods is hardly to recover surface machining error, even in coarse scale.

To verify the performance of proposed approach on practical machined work piece, we machine a Peaks surface and use CMM to get the stand machining errors with 6561 samples. Then we use $r=1/16$ sparse samples to reconstruct the final results under sparse sampling.
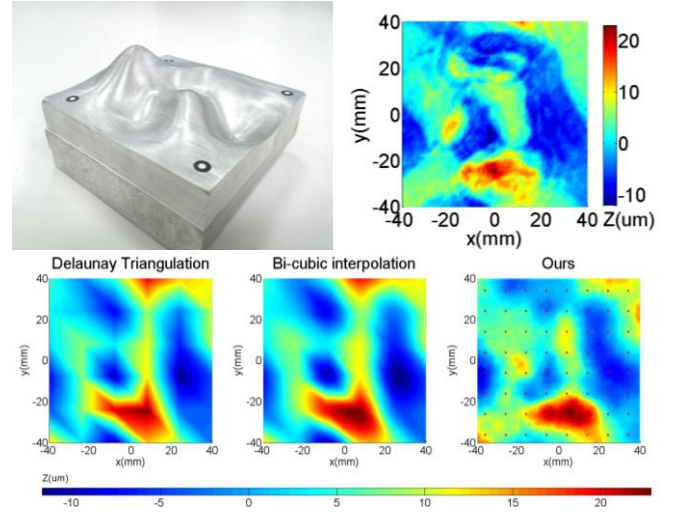
Fig.10. Machined Peaks workpiece and benchmark machining errors, measured by Hexagon Leitz PMM-Xi Coordinates Measuring Machine. In $r=16$, our approach could achieve more accurate results. The red points on ours results show samples.

Experiment shows proposed method could reconstruct better results than competitors without tuning. Results' smoothness reflects the continuity of the machining process. It worth to note that there is still room for improvement by introducing training data which is similar to the measurement target.

IV.   CONCLUSION

This paper proposes a novel measurement strategy to improve the machining error evaluation efficiency of CMM. We introduce the down-sampled method to reduce the number of samples and consumed time simultaneously. Then, we propose an encoder-decoder based deep learning model to reconstruct dense surface error, with sparse measurement inputs. Besides, to prepare datasets for model training, this paper adopts fractional Brownian motion to describe surface machining errors, which could be used to synthesis abundant data. With synthesis datasets and sophisticated architecture, proposed model could learn to reconstruct dense machining error from sparse measurement. The reconstruction process also could be accelerated by high performance computing, which could further improve measurement efficiency. In conclusion, this paper proposes a novel perspective for efficiently accurate measurement, which has theoretical and engineering significance for precision measurement, precision processing, and high-end equipment manufacturing.

Future work will extend the proposed strategy into more extensive measurement fields, and strengthen model generalization by introducing more accurate mathematical model and practical measurement dataset. Besides, it is worth trying advanced deep learning models, such as generative adversarial network, variational auto-encoder, transformers and Bayes networks, which could potentially improve reconstruction accuracy under this framework.

<center>REFERENCES</center>

[1] J. Lai, J. Fu, H. Shen, W. Gan, and Z. Chen, "Machining error inspection of T-spline surface by on-machine measurement," International Journal of Precision Engineering and Manufacturing, vol. 16, pp. 433-439, 2015.

[2] L. Sun, M. Ren and Y. Yin, "Domain-specific Gaussian process-based intelligent sampling for inspection planning of complex surfaces," International Journal of Production Research, vol. 55, pp. 5564-5578, 2017.

[3] L. Sun "Research on Complex Optical Surface Reconstruction and Multi-Sensor Data Fusion Method based on Gaussian Process," Doctor Thesis, Shanghai Jiao Tong University, Shanghai, P.R. China, 2018.

[4] G. Trieb, K. Schepperle and K. Aubele, "Control for coordinate measuring instruments,": Google Patents, 1988.

[5] X. D. Zhang, K. Pu, and Y. W. Dong, "Optimal sampling strategy for aero-engine blade inspection with coordinate measuring machine," Journal of Aerospace Power, vol. 34, pp. 168-176, 2019.

[6] J. Wang, X. Jiang, L. A. Blunt, R. K. Leach, and P. J. Scott, "Efficiency of adaptive sampling in surface texture measurement for structured surfaces," in Journal of Physics: Conference Series, 2011, p. 012017.

[7] P. Dierckx, Curve and surface fitting with splines: Oxford University Press, 1995.

[8] W. Wang, H. Pottmann and Y. Liu, "Fitting B-spline curves to point clouds by curvature-based squared distance minimization," ACM Transactions on Graphics (ToG), vol. 25, pp. 214-238, 2006.

[9] S. Y. Gatilov, "Vectorizing NURBS surface evaluation with basis functions in power basis," Computer-Aided Design, vol. 73, pp. 26-35, 2016.

[10] S. Fortune, "Voronoi diagrams and Delaunay triangulations," in Computing in Euclidean geometry: World Scientific, 1995, pp. 225-265.

[11] D. OuYang and H. Feng, "Reconstruction of 2D polygonal curves and 3D triangular surfaces via clustering of Delaunay circles/spheres," Computer-aided design, vol. 43, pp. 839-847, 2011.

[12] H. Q. Dinh, G. Turk and G. Slabaugh, "Reconstructing surfaces using anisotropic basis functions," in Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, 2001, pp. 606-613.

[13] S. Ferrari, I. Frosio, V. Piuri, and N. A. Borghese, "Automatic multiscale meshing through HRBF networks," IEEE Transactions on Instrumentation and Measurement, vol. 54, pp. 1463-1470, 2005.

[14] L. J. Sun, M. J. Ren and Y. H. Yin, "Gaussian process based intelligent sampling for measuring nano-structure surfaces," in 8th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Optical Test, Measurement Technology, and Equipment, 2016, p. 96840L.

[15] W. Zhou, M. Ren, Y. Tao, L. Sun, and L. Zhu, "Enhancing the metrological performance of non-raster scanning probe microscopy using Gaussian process regression," Measurement Science and Technology, vol. 30, p. 095004, 2019.

[16] C. K. Williams and C. E. Rasmussen, Gaussian processes for machine learning vol. 2: MIT press Cambridge, MA, 2006.

[17] I. Goodfellow, Y. Bengio and A. Courville, Deep learning: MIT press, 2016.

[18] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," IEEE transactions on pattern analysis and machine intelligence, vol. 38, pp. 295-307, 2015.

[19] C. Dong, C. C. Loy and X. Tang, "Accelerating the super-resolution convolutional neural network," in European conference on computer vision, 2016, pp. 391-407.

[20] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1874-1883.

[21] J. Kim, J. Kwon Lee and K. Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1646-1654.

[22] Kim J, Kwon Lee J, Mu Lee K. 2016. Deeply-recursive convolutional network for image super-resolution[C]//1637-1645.

[23] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, and Z. Wang, "Photo-realistic single image super-resolution using a generative adversarial network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4681-4690.

[24] Y. Tai, J. Yang and X. Liu, "Image super-resolution via deep recursive residual network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 3147-3155.

[25] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2017, pp. 136-144.

[26] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 4799-4807.

[27] S. Anwar, S. Khan and N. Barnes, "A Deep Journey into Super-resolution: A Survey," ACM Computing Surveys (CSUR), vol. 53, pp. 1-34, 2020.

[28] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang, Stochastic calculus for fractional Brownian motion and applications: Springer Science & Business Media, 2008.

[29] M. J. Ren, C. F. Cheung and L. B. Kong, "A task specific uncertainty analysis method for least-squares-based form characterization of ultra-precision freeform surfaces," Measurement Science and Technology, vol. 23, p. 054005, 2012.

[30] W. Lee, Y. Lee and C. Wei, "Machining Error Compensation Based on the Reconstructed Free-Form Surfaces of CAD Models," in 2018 International Conference on System Science and Engineering (ICSSE), 2018, pp. 1-4.

<center>447</center>

Architecture and Implementation of IoT Middleware for Ground Support Systems in Launch Site

Litian Xiao*
Beijing Special Engineering Design and Research Institute
Beijing, China
xiao_litian@sina.com

Nan Xiao
China Aerospace Academy of System Science and Engineering
Beijing, China
cloudia@126.com

Mengyuan Li
Beijing Special Engineering Design and Research Institute
Beijing, China
lmy2000@sohu.com

Shanshan Xie
Information Support Lab, Equipment Department, Aerospace System Ministry
Beijing, China
shanshan_xie2020@126.com

Kewen Hou
Beijing Special Engineering Design and Research Institute
Beijing, China
houkewen@126.com

Yuliang Li
Beijing Special Engineering Design and Research Institute
Beijing, China
liyuliang1980@yeah.net

Abstract—The ground support systems (GSS) at launch sites include the test and launch system of launch vehicles, satellites, payloads, test equipment, and ground support facilities such as propellant loading, gas supply, etc. Various devices or equipment in GSS are interconnected through Internet of Things (IoT) to achieve integrated application, e.g. data interaction, interoperation, and application management. Because there are a large number of heterogeneous software and hardware in GSS, an IoT middleware is the key to the integrated application of GSS. We propose an architecture of IoT middleware for GSS at the launch site. This architecture solves the heterogeneity problems of GSS in launch sites such as heterogeneity between embedded sensor devices and mobile terminals, conversion between mobile network and backbone network, communication and data exchange, integration of different application systems, and system interconnection. Through the integrated portal of IoT middleware based on the architecture, heterogeneous systems of ground support facilities are realized on integrated application and deployment. An integrated access gateway is developed to make the IoT middleware accessed various devices or equipment of GSS. The integrated access gateway has good extensibility and adaptability to accessing interfaces or connectors. A simple application case of the IoT middleware is introduced in the paper. Good results have been achieved through practical application in China launch sites.

Keywords—IoT middleware, architecture and implementation, ground support systems, test and launch, launch site

I. INTRODUCTION

The ground support systems (GSS) involve the test and launch equipment and ground support facilities for launch vehicles, satellites, payloads, flight products, propellant loading, gas supply, air conditioning, etc. Now the equipment and the facilities have been upgraded by IoT technology for integrally accessing GSS. Thus GSS can realize the integrated application in launch sites and achieve information sharing, centralized facilities management and comprehensive applications such as test and launch system, automatic control system (on gas supply, propellant loading, and bridge crane), enclosed area environmental monitoring and control system, power quality monitoring and control system, health monitoring system of the launch tower structure, etc. [1] The applications also include spacecraft safety monitoring and support, ground facilities management system, personnel distribution management system, safety precaution system,

full-process dynamic management system for test and launch, etc.[2]

Because IoT application in launch sites faces a series of problems brought about by a large number of applications, the variety of equipment, and the heterogeneity of hardware and software in GSS. It leads to complex access, poor interoperability, delayed data interaction, and difficult application management and integration. We propose the architecture of IoT middleware for GSS at launch sites (IoT-MLS). IoT-MLS solves the heterogeneity application problems of GSS in launch sites.

IoT-MLS is responsible for shielding the heterogeneity of the hardware devices, network and service platform for the lower hardware. It makes the devices and equipment of launch site to be unified access and centralized management. It supports application development, runtime sharing, and open interoperability for the upper application. And it guarantees the reliable deployment and management of IoT applications for GSS. IoT-MLS is the core and the foundation of IoT construction at the launch site. IoT-MLS based on the proposed architecture has been developed, deployed, and achieved good results in China launch sites.

II. RELATED WORKS AND REQUIREMENTS

A. IoT Middleware

IoT-MLS is the key infrastructure and common platform of IoT applications in launch sites. It integrates all operation applications to achieve the comprehensive application of resource sharing, data fusion, and intelligent decision-making between different platforms in GSS [3] [4] [5]. The IoT-MLS hierarchy diagram is shown in Fig.1 where MT is a mobile terminal and NFC is near field communication.

The conventional IoT middleware can be divided into application service middleware, embedded middleware, and hybrid middleware.

The application service middleware is usually deployed on the application server based on TCP/IP networks. It communicates with different devices by loading drivers or adapters, and no additional hardware is required. The server is directly connected to the devices. The middleware converges the data of different devices into the event data needed by IoT applications. Typical application service middleware includes Oracle RFID middleware products [6][7] and Rifidi [8][9] simulation development platform which is open-source. Its main advantage is easy to deploy. The disadvantages are the

---

two sides. One side is hardly conducive to distributed deployment because of the direct connection between the servers and the devices. Another side is easy to single point failure because the data is centrally processed by the servers. A single point failure is a very serious problem in GSS at the launch site.
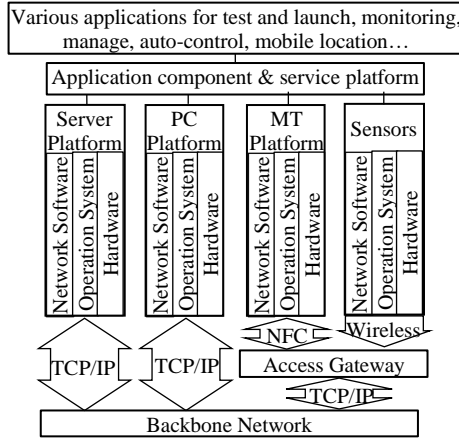
Fig. 1. IoT-MLS hierarchy diagram.

The embedded middleware makes use of devices to quickly filter and convert data or other preprocessing. Data processing is close to the source of the data so that network traffic and the stress of the application server can be reduced. The embedded middleware mainly is wireless sensor network (WSN) middleware such as TinyDB [10] [11] and MetaQ [12]. The main advantages of embedded middleware are that sensor nodes bear data processing capacity and reduce the stress on networks and servers. The main drawbacks are limited services in which the middleware only supports simple services, e.g. node discovery, data filtering and processing, and code migration. The middleware also does not support heavyweight computing and storage services.

The hybrid middleware divides the application service middleware into two parts. The one part is the common service and runtime support environment. Another part is the device abstraction layer and runtime support environment. The former is deployed on the server. The latter is deployed in a special access gateway. The more representative hybrid middleware is the IBM RFID middleware products [13] [14] and the SYBASE RFID middleware products [15] [16]. Its access gateway is a bottleneck on accessing devices because different access gateways need to be developed for devices.

The conventional application service middleware based on server platform has been developed to provide data storage, numerical calculation, statistics mining, policy control, TCP/IP communication services, and other common services [17] [18] [19]. Because the features of sensor nodes are mass heterogeneous data transmission, the storage and forwarding functions including fusing data and computing are deployed in embedded middleware to reduce the cost of application server and backbone network [20]. The cost of accessed devices is enhanced for developing additional embedded middleware.

The tight coupling service mode of conventional IoT middleware exists poor extensibility, limited capacity of data processing and device access, insufficient service dynamic matching optimization, and complex middleware construction. For solving these problems, microservice architecture combined with middleware is proposed and gains expandability. The microservice can be loaded by protocol plug-in and shield heterogeneous communication devices. [21] Although microservice-based middleware has the characteristics of rapid deployment, scalability, and shielding against communication heterogeneity [22][23], it also has insufficient adaptability and transforming real-time of middleware access interfaces as well as inapplicable to control system for facilities.

B. Requirements at Launch Sites

In launch sites, the heterogeneity of embedded sensor devices, ground support facilities, and mobile terminals are not only reflected in the hardware devices, operating systems, and network protocols produced by different manufacturers, but also in the storage, computing, and communication of the devices. As various intelligent terminal devices access (or exit) the IoT of the launch site, or move locations, all may cause changes in the network topology. The IoT middleware must be able to solve the changes in mobility and network environment.

The data types and data access control ways of the wired backbone network and the wireless sensor access network are different in the launch site. The network communication service of the IoT middleware needs to combine multiple network communication modes according to different application requirements. Therefore, the communication mechanism needs to be designed to support the exchange of multiple types of access and data.

The conventional middleware of server-based application service [7] has developed relatively maturely, and a large number of launch site equipment has caused huge network overhead. The situation influences the efficiency of IoT applications, especially for unattended applications with high real-time requirements. Part of the data fusion computing and store-and-forward functions need to reduce the overhead of the application server and backbone network. A large number of devices in ground support facilities need to have the heterogeneous interfaces and real-time requirements on IoT middleware.

There are many types of nodes in GSS at launch sites. Some nodes have different operating systems, and some do not even have an operating system. How to transplant these nodes into middleware need be considered for the integrated operation of IoT middleware.

III. Architecture Design of IoT-MLS

The IoT-MLS of GSS should be able to reduce the cost of operation and maintenance and improve the safety and emergency rescue capabilities of ground facilities and equipment. It is designed by the above requirements and consisted of seven parts including basic software and hardware environments, databases, common services, application components, integrated portals, developing tools, standards, and specifications, etc., as shown in Fig.2.

1) Basic software and hardware environment: It mainly includes hardware and operation systems such as networks, computers, servers, and software environments such as database management systems and office software. It is the foundation and supports operation environments for middleware. It also provides a redundant environment for application services and heavyweight computing for the devices without or with insufficient computing capabilities.
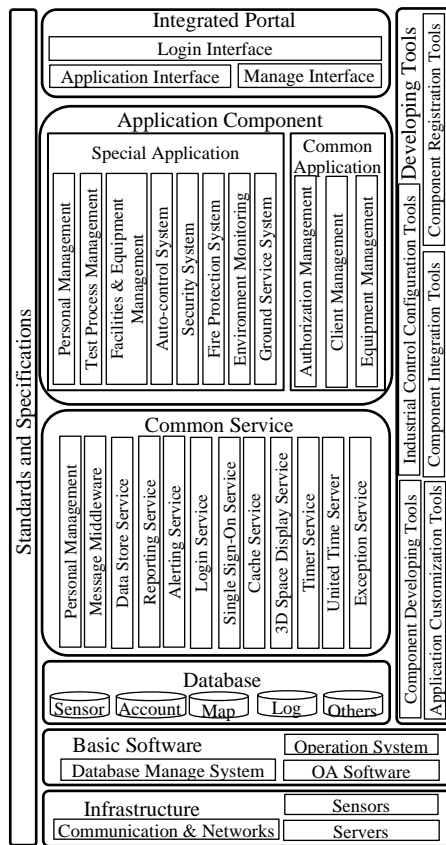
Fig. 2. IoT-MLS architecture.

2) Database: The database is composed of a sensor database, account database, map database, and log database, and so on. It supplies the data service of IoT applications and enhances storage capacities for required devices or equipment.

3) Common service: Common service consists of personal management, message middleware, data store service, alerting service, login service, 3D space display service, united time service, and so on. The same and similar services in IoT applications are concentrated in IoT-MLS to provide unified customization services for GSS. It lightens the performance loads of GSS.

4) Application components: Special applications are the developed components that perform the specific functions of IoT applications. The components of the special application include test process management, facilities and equipment management, auto-control system, security system, fire protection system, environment monitoring, ground service system, and other extensible applications. To improve common management for component application, the common application provides authorization management, client management, equipment management, and other public management.

The operating environment of IoT-MLS provides the operating container for the components, which enables the IoT application component to work in the IoT-MLS platform. IoT-MLS provides various services for component operation support on events, data, network, protocol-driven, and data exchange.

5) Integrated portal: It provides users with a unified login interface. Component developers and managers use the management interface to login in IoT-MLS. Users use the application interface to apply various functions and components in IoT-MLS. User can customize their own business applications through login in the integrated portal. It is integrated with the operation platform and environment.

6) Developing tools: It provides users to develop specialized components and IoT applications. The developing tools of IoT-MLS provide component registration tools, component integration tools, application customization tools, industrial control configuration tools, and component developing tools. Through the tools, users can flexibly customize and combine the developed component into specific applications. It integrally completes business, process, and user functions on IoT-MLS.

7) Standards and specifications: It specifies standards and specifications used by IoT-MLS on protocols, interface, development, integration, application, etc.

IoT-MLS architecture is suitable for the heterogeneity of embedded sensor devices, ground facilities, and mobile terminals because of hardware devices, operating systems, and network protocols produced by different manufacturers. It makes the applications of GSS less coupled to specific devices and equipment. It can coordinate the capacity differences and support on storage, computing, and communication.

Because the infrastructure can be increased in servers and customized in application components, the architecture makes the network and the application environment keep the stability and real-time when various terminal devices or mobile equipment access (or exit) IoT at the launch site. The data types and data access control methods are different in the wired backbone network and the wireless sensor access network. The network communication service combines multiple communication modes based on different application requirements. Forming a set of communication and information exchange standards under the architecture, a communication mechanism is designed to support the access and exchange of various types of data within facilities. It is the developed foundation of access gateway.

All application components in the IoT of launch sites can be customized through the auxiliary tools of IoT-MLS. After each component is developed, a specific business can be independently operated and completed. IoT-MLS can be integrated through the development tools to complete the data exchange among the individual components. Thus, the business applications are not only independent of each other in their business functions, but also interconnected in the data flow. In this way, IoT in the launch site can be constructed by blocks. The approach can reduce the coupling within IoT-MLS and improve the system's expansibility and maintainability.

IV. IMPLEMENT OF ACCESS GATEWAY AND I/O INTEGRATED CONVERSION

A. Integrated Access Gateway

In order to meet the various interface requirements of the ground facilities, we design and develop a modular integrated

access gateway with plug and play. The integrated access gateway (IAG) is mainly composed of an operation panel, a motherboard, and a pluggable I/O daughter card.

Users can realize the configuration management of the gateway through touch control of the operation panel. The operation panel can display the monitoring data of the sensors belonging to the operation, configure the network node equipment. IAG supports multi-user access and has a multi-layer password authentication mechanism. IAG is enabled independently to process data and support distributed access applications. IAG motherboard is a core and completes the communication protocol conversion and coordinates the data transmission of multiple I/O daughtercard which is used to provide multiple types of communication signal processing. The I/O daughter cards have developed for ZigBee, RFID, UWB, RS485, RJ45, RS232, 4G/5G, audio, and video, etc.

The architecture of the FPGA-based IAG motherboard is shown in Fig. 3. The IAG motherboard is designed to three parts: general control block (GCB), backplane slot, and rear transport block (RTB). FPGA is used for logic and algorithm processing. The base programmable logic module (BPLM) of GCB is used for general and standard access protocols processing. BPLM mainly deals with fixed I/O management, and conversion algorithms and logic. Applying dedicated programmable logic module (ADPLM) provides users with programmable applications to facilitate different access extensions and I/O data conversions. Especially ADPLM can be suitable for new access equipment at the launch site, e.g. the different satellite test equipment change greatly. IAG is only refreshed in ADPLM and matched the appropriate daughter cards. It makes IAG more adaptable [24].

FPGA-Based Motherboard
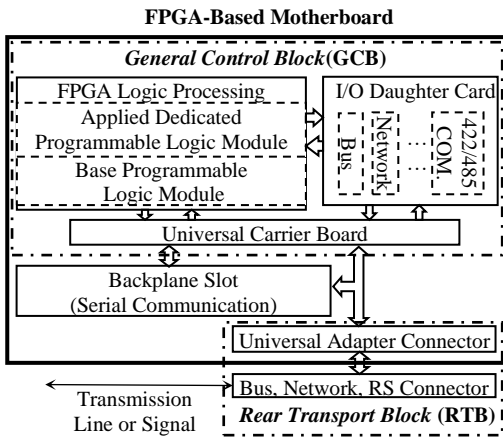
Fig. 3.   The architecture of the FPGA-based IAG motherboard.

A universal carrier board is a slot carrying motherboard and I/O daughter cards. I/O daughter cards are reserved on the IAG motherboard for processing standard or customized I/O signals. the FPGA core processing module processes protocol conversion, I/O reading, and writing through the universal carrier board. For example, if you need to add a 5G I/O application, you only need to customize a 5G daughtercard, refresh the ADPLM, and IAG can directly apply the 5G communication mode. The motherboard is also inserted into the backplane slot of the chassis through the universal carrier board, which can be used for stacking. Different I/O customization requirements can be realized by configuring different daughter cards.

Rear Transport Block (RTB) is an interface connector for communication signals between the GCB and the connector

interface on a chassis. These connectors reserve the existing universal communication interface. The designed conversion standard interface can also access the new customized communication interface connector. This system makes IAG highly adaptable and extensible.

B. Customized I/O for IAG

In customized I/O daughtercard of IAG, I/O communication information needs to generate fixed content and fixed length. Because this information has no length limit, it can contain more data than routing information. Each message still contains fixed/predefined message content. However, if the message length is too long, piecewise multiple messages will be sent. Each of piecewise multiple messages will contain a fixed/predefined message content.

Under customized I/O, the point-to-point information with the external access system will be generated by the communication interface in the specified format. The communication protocol will be defined according to the communication connection requirements between IAG and the external access system.

To ensure that the user-defined I/O application is as simple as possible, a predefined information format with fixed data content will be used. The communication daughter card provides electrical isolation between the security system and the non-security system. As a communication processing device, functional isolation and data isolation can be performed. Data isolation will be ensured by the management constraints in IAG application guidelines.
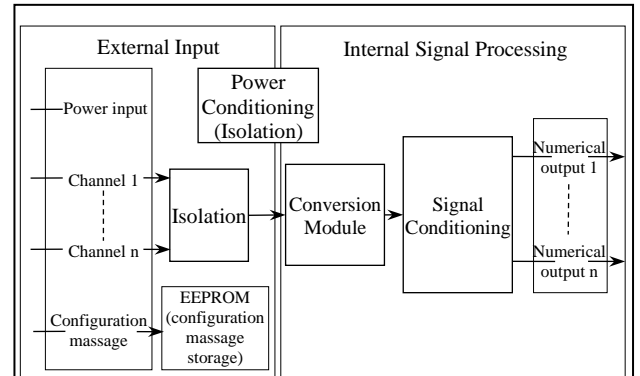
Fig. 4.   Schematic diagram of input signal decoding.

The function of the I/O daughtercard is to convert various protocols, and signal encoding and decoding. Fig.4 shows I/O daughtercard that the input signals are decoded by the conversion module on the I/O daughter card according to the protocol. Conversely, the encoding output is similar. The customized I/O daughter cards within IAG ensure adaptability and transforming real-time of middleware access interfaces.

V.   IoT-MLS Construction Effect and Application Cases

A. IoT-MLS Construction Effect

IoT-MLS provides an integrated platform for IoT development and application in launch sites. In GSS application, developers log in to IoT-MLS through the integrated portal to customize the application calling common services and common applications. The developing tools supply the development of specialized applications. Developed special applications can also be called. The

customized application is developed and will be published by developing tools and the registration management mechanism. According to standard protocols, the underlying GSS hardware accessing is managed by IAG and IoT-MLS. Users log in to the integrated portal to access IoT-MLS and use the corresponding functions. IoT-MLS solves the problems of sharing applications across heterogeneous facilities and equipment, which realizes heterogeneous data interaction. IoT-MLS provides an application platform with strong interoperability and equipment independence. This is necessary to implement the application of distributed remote GSS.

Through GSS construction, IoT in launch sites can be realized the integration and interconnection of broadband wireless network, wired IP network, sensor network, transmission network, and other networks within the unified architecture. Various sensor equipment and GSS equipment are accessed into IoT-MLS by IAG for omnidirectional sensing and data collection. IoT-MLS is a basic application platform for GSS sensing and data processing by means of diversified unified access and data protocols of I/O interface. The common application and services combined with 3D visualization present comprehensive situational awareness.

The ground facilities and equipment realize the automatic remote execution capability through IoT-MLS, which supports the requirement of remote control and unattended operation in launch sites. IoT-MLS, as a basic platform and infrastructure, supports the construction application of IoT, big data, cloud computing, knowledge bases, 3D visual presentation, etc. It provides means to realize the unified health and life-cycle management for ground facilities and equipment of all major systems in GSS such as the launch station, non-standard equipment, propellant loading, air conditioning, power supply and distribution, and fire protection. The application also provides good flexibility, compatibility and extensibility for the continuous development and dynamic variable demand in launch sites. The following is a simple case to illustrate the process of development and application.

B. Application Development

The application mode of IoT-MLS adopts the trinity mode of custom-integration-operation. A special application is custom-developed by component developing tools. It is integrated the applications and is built into IoT-MLS by component integration tool. The integrated portal enables users to develop, customize, and apply IoT-MLS. The application mode is shown in Fig.5.
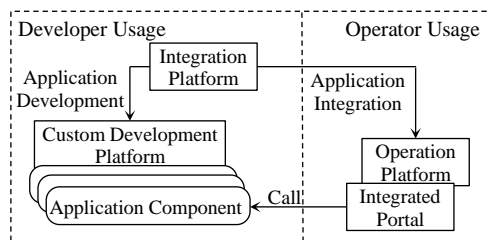
Fig. 5.   Mode diagram of IoT-MLS application.

The developers or users of the customized development and the integration construct the specific system and applications on IoT-MLS for operators/end-users. Operators/end-users use and perform specific system operations by means of IoT-MLS integrated portal and operation platform. The specific system construction and working principle diagram are shown in Fig.6. After the development of each component is completed in IoT-MLS, IoT-MLS can run independently and complete a specific business for user applications.
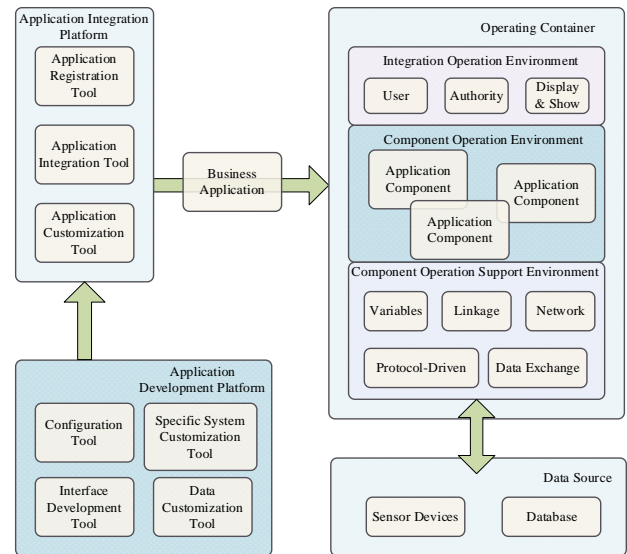
Fig. 6.   Specific system construction and working principle diagram.

Fig.6 illustrates that users can flexibly combine the developed applications and customize the data flow interaction relationship between applications. Thus users can complete the establishment of application services, processes, and user functions. IoT-MLS integrates business applications, user rights, and data interaction rules by loading the configuration of user customization, which provides the user's final operation functions.

C. Application Case

As an example, a simple case introduces the application flow of the IoT-MLS development and deployment for personnel location.

The case is that the position of the personnel in the room is displayed in real-time. It utilizes the ZigBee anchor node and positioning tag card to realize the personnel location located in the room. Its positioning principle is described as following.

The anchor node is fixed with the room coordinate. The positioning tag card carried by personnel sends the wireless signal to the anchor node in real-time. And the anchor node sends the received signal to the location service component. The component converts the field strength of the received signal to get the distance between the positioning tag card and each anchor node. Then it calculates the coordinates of the positioning tag card by using the trilateral measurement algorithm combined with the coordinate of the anchor node. Main sensing devices are 3 ZigBee anchor nodes and 4 positioning tag cards. The personnel location application has 6 steps for development and deployment.
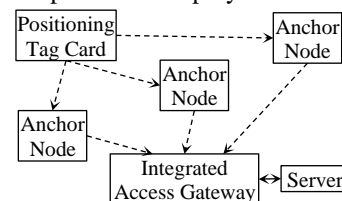
Fig. 7.   Connection relation of equipment hardware.

1) Device connection: As shown in Fig. 7, three ZigBee nodes are wirelessly connected to a positioning tag card and IAG via ZigBee. IAG connects to the server via a network cable.

2) Driver development: The protocol driver component of the ZigBee anchor node is developed by component development tool and loaded into IAG.

3) Service component development: The positioning service components are developed by component development tools and loaded them into IAG.

4) Positioning logic design: As shown in Fig. 8, the data of the Zigbee anchor node is abstracted into a field strength data component. It connects the output of the field strength data component with the input of the location service component. The output of the location service component is connected to the computing location result component. Then the logical design of the location application is completed, and the location coordinate is generated by the sensing data through the flow direction from left to right.
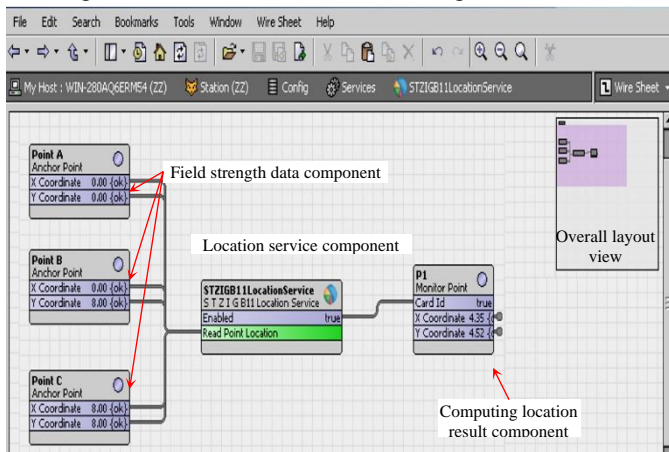
Fig. 8. Graphic programming.

5) Interface design: As shown in Fig. 9, anchor node icons (labeled as star icons)and positioning tag card icons (labeled as circle icon 1 to 4) are added into the display interface labeled on the floor of the room plan by interface development tool. This completes the application development.

6) Application deployment: The computing location result component is associated with the positioning tag card icons. The positions of the circle icons in Fig. 9 are changed with the coordinates change of positioning tag card. So far, application deployment is completed.

VI. CONCLUSION

The research is based on the requirements of IoT application in the launch site. IoT-MLS architecture is proposed in the paper. The implementation scheme combined hardware and software are designed for IoT-MLS and IAG. A simple case of personnel location application is implemented by graphical programming, which demonstrates the IoT-MLS development and deployment process. Now IoT-MLS is being progressively deployed at China launch sites [25]. IoT-MLS effectively solves the shortcomings of existing and conventional middleware products in the IoT application of

launch sites. For the complex equipment access of GSS, the implementation scheme achieves the requirements on protocol processing real-time, access flexibility, IoT adaptability, and application scalability.
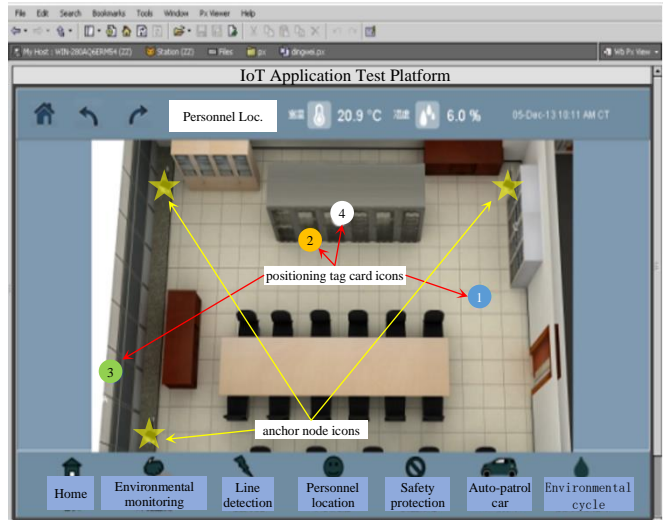
Fig. 9. The application interface of personnel location in a room.

Currently, the type and number of components developed in the research cannot fully cover the IoT application requirements of GSS at launch sites. The next step needs to strengthen the development of components and increase more common services so that IoT application efficiency can be improved further.

For GSS in launch sites, IoT-MLS application lays a good foundation to realize completely unattended and intelligent operation, the applications of big data and cloud platform, and intelligent maintenance management based on machine learning [26][27]. These applications can improve launch efficiency and success rate of launch mission.

REFERENCES

[1] Litian Xiao, *et al*."A Hierarchical Framework for Intelligent Launch Site System", 2018 IEEE International Conference on Information and Automation (ICIA), Wuyishan, China, Aug.20-25, 2018.

[2] Litian Xiao, *et al*., "System Architecture and Construction Approach for Intelligent Space Launch Site", Journal of Transactions on Advances in Intelligent System and Computing, 2018, vol.865, pp.397-404.

[3] S. Huda, A. Wesam, E.S. Khair, "Internet of Things: A Review to Support IoT Architecture's Design," IT-DREPS Conference, Amman, Jordan, Dec 6-8, 2017.

[4] S. Li, L.D. Xu, S. Zhao, "The internet of things: a survey," Information Systems Frontiers, 2015, vol.17(2), pp.243-259.

[5] J. Lin, W. Yu, N. Zhang, X. Yang, *et al*., "A survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and applications," IEEE Internet of Things Journal, 2017, vol. 99, pp.1-11.

[6] Oracle, http://www.oracle.com, 2018.

[7] N Deng, "RFID Technology and Network Construction in the Internet of Things," International Conference on Computer Science & Service System, 2012, pp.979-982.

[8] Rifidi, http://transcends.co, 2018.

[9] Ibrahim Mezzah, *et al*., "Emulation-based fault analysis on RFID tags for robustness and security evaluation", Microelectronics Reliability, 2017, vol. 69, pp.115-125.

[10] TinyDB, http://tinydb.readthedocs.org, 2018.

[11] Dong-Oh Kim, *et al*., "Spatial TinyDB: A Spatial Sensor Database System for the USN Environment", International Journal of Distributed Sensor Networks, 2013, vol.2013, pp.1-10.

[12] J. Liao, X. Zhuang, R. Fan, X. Peng, "Toward a General Distributed Messaging Framework for Online Transaction Processing Applications," IEEE Access, 2017, vol. 99, pp.1-11.

[13] IBM, http://www.ibm.com, 2018.

[14] A.M. Khan, A. Khaparde, V.P. Savanur, "Self-aware inventory system based on RFID, sensors and IBM security directory integrator," International Conference on Inventive Computation and Technologies, 2017, pp.1-4.

[15] SYBASE, http://infocenter.sybase.com, 2018.

[16] Siti Salwani Yaacob, *et al*., "A Review on Complex Event Processing in RFID System", International Journal on Information Technology, 2018, vol. 8(2018), pp.1154-1163.

[17] V. Gazis, M. Görtz, M. Huber, *et al*., "A survey of technologies for the internet of things," International Wireless Communications & Mobile Computing Conference, USA, Dec. 2015.

[18] M. Jose, "P2.42: Latest Development in Advanced Sensors at Kennedy Space Center (KSC)," IEEE Sensors, 2002, vol. 2(2), pp.1728-1733.

[19] Giacomo Benincasa, *et al*., "Agile Communication Middleware for Next-Generation Mobile Heterogeneous Networks", IEEE Software, 2014, vol. 31 (2), pp.54-61.

[20] Ali Alfoudi, *et al*., "Slicing Architecture for Managing User Mobility in Next-Generation Heterogeneous Wireless Networks", IEEE Network, 2020, vol. 2019, pp.1-4.

[21] Björn Butzin, *et al*. "Microservices approach for the internet of things ", 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), Berlin, Germany, Sept.6-9, 2016.

[22] Datta Soumya Kanti, *et al*. "IoT and microservices-based testbed for connected car services", 5th Workshop on Smart Vehicles: Connectivity Technologies and ITS Applications, Chania, Greece, June 12-15, 2018.

[23] Cleber Santana, *et al*. "A reliable architecture based on reactive microservices for IoT applications", Proceedings of the 25th Brazillian Symposium on Multimedia and the Web, Rio de Janeiro, Brazil, Oct. 15-19, 2019.

[24] Litian Xiao, *et al*." An Architecture of FPGA-Based Controller on Automatic Control System for Test and Launch in Launch Site ", ACM 2nd Int. Conf. on Computer Science and Application Engineering, Hohhot, China, Aug.5-7, 2018.

[25] Litian Xiao, *et al*." IoT-Based Architecture of Intelligent Test and Launch Control System in Advanced Launch Site ", 2019 70th International Astronautical Congress (IAC), Washington D.C., USA, Oct.20-25, 2019.

[26] Litian Xiao, *et al*." Intelligent Architecture and Hybrid Model of Ground and Launch System for Advanced Launch Site ", 2019 IEEE AeroSpace Conference, Bigsky, USA, Mar.5-9, 2019.

[27] Dong Wei, *et al*." Application Research of Big Data for Launch Support System at Space Launch Site ", 3rd Int. Conf. on Computer Science and Application Engineering, Sanya, China, Oct.25-27, 2019.

Dual-lane Phononic Crystal for Low-frequency Elastic Wave Attenuation

Jiawen Xu*
Instrument Science and Engineering
Southeast University
Nanjing, China
jiawen.xu@seu.edu.cn

Ruqiang Yan
Mechanical Engineering
Xi'an Jiaotong University
Xi'an, China
yanruqiang@xjtu.edu.cn

Abstract— We report a phononic crystal system for achieving ultra-broadband elastic wave attenuation by taking advantages of self-cancellation. This concept is applied to a finitely long dual-lane phononic crystal beam. The two lanes are designed to split the incident elastic wave into two parts and introduce a phase difference between the waves that yields significant wave attenuation. The simulation result in this paper show that the wave attenuation capacity of proposed dual-lane phononic crystal system is remarkably improved.

Keywords—Phononic crystal, Dispersion relation modulation, Destructive interference, Elastic wave attenuation

I. INTRODUCTION

Phononic crystals as well as metamaterials, which are artificially designed periodic structures, have attracted intensive attention due to their promising advantages in wave manipulation and attenuation [1-5]. PCs and metamaterials have applications of negative refraction [6], acoustic cloaking [4], wave focusing [7,8], wave attenuation [3,9] and vibration mode tailoring [10]. Moreover, dispersion curves of PCs and metamaterials may exhibit a negative slope, yielding negative refraction phenomenon [6]. The unique features of phononic crystals and metamaterials are originated from the mechanisms of Bragg scattering and local resonance, respectively [2,11,12]. The Bragg scattering phenomenon occurs when the elastic wave-length is at the same level to the dimension of a unit-cell [13,14]. Consider the fact that the wave-length of an elastic wave is roughly inversely proportional to its frequency, the application of PCs in the low-frequency regime requires large-size unit-cells. On the other hand, metamaterials take advantages of the local resonance behavior that does not rely on the unit-cell dimension possess extraordinary capability of manipulating low-frequency elastic waves at sub-wavelength scales [1,15].

Extensive efforts have been devoted to attenuating elastic wave using PCs and metamaterials. For example, a low-frequency bandgap would be generated around the resonating frequency of the identify resonators integrated in the structures periodically. The generated bandgap in such a kind of metamaterials exhibited a strong vibration attenuation effect. Explorations were carried out for the vibration control in structural bars and beams [5,14]. The width of the of was demonstrated to be expandable by parametric optimization [20]. Normally, the mechanical metamaterials and PCs have fixed bandgap behavior [5]. To introduce online tunability, piezoelectric transducers have been adopted due the electro-mechanical coupling [9,19]. A negative capacitance element, a non-resonating shunt circuit, was integrated in piezoelectric metamaterials for the purpose of reducing the stiffness, thereby modifying the bandgap features through the modifying the value of the negative capacitance [17].

Alternatively, piezoelectric metamaterial with unit-cell level LC resonating can produce a bandgap. For example, periodic piezoelectric transducers with inductance shunt circuits are adopted for wave attenuation [20]. Piezoelectric phononic crystals and metamaterials with shunt circuits have advantages over the mechanical one due to their simple configuration and adaptivity. The external shunt circuits allow adjusting of bandgap towards a desired frequency range without modifying the mechanical designs of the system.

On the other hand, the application of PCs and metamaterials in the field of wave attenuation is largely restricted by their narrow bandwidths. To overcome this limitation, efforts have been devoted into achieving bandgap enlargement. It was demonstrated that the mechanical connection between adjunct unit-cells can effectively enlarge the bandgap width of the elastic metamaterial [19]. Moreover, tunable wave attenuation is highlighted in the piezoelectric phononic crystal/ metamaterial for applications at wide frequency range. Unit-cell level optimization of a piezoelectric metamaterial was performed and it was demonstrated that the enhancement of electromechanical coupling could widen the bandgap width [20]. More recently, it was illustrated that incorporating nonlinearities into the local resonators of elastic metamaterials is an alternative strategy to broaden the bandgap [21-22].

In this research, we propose a PC plate with coupled lanes that exhibits promising wave attenuation capability. The key of the proposed system lies in producing different phase shifts of elastic waves in the separated lanes. A proper tuning of the phase difference results in the destructive interference phenomenon that yields enhanced wave attenuation in aspects of bandwidth and depth. The rest of this paper is organized as follows. In Section 2, our novel concept of dual-lane is outlined first. In Section 3, governing equation of a unit-cell is derived and a model is presented for evaluation of the unit-cell's dispersion relation. In section 3, phase shift characteristics of the lanes are investigated. Section 4 provides correlated analysis for validation and insights of the wave attenuation features of the PC with coupled lanes. Section 5 presents the concluding remarks.

II. CONCEPTUAL DEMONSTRATION

The proposed PC plate that consists of a series of coupled PC lanes is outline in Fig. 1. Unlike traditional systems based on local resonance or Bragg scattering mechanisms, the proposed system takes advantages of destructive interference stemming from the phase shift difference between the coupled lanes. Destructive interference has been widely adopted in the design of antenna for directional wave minimization. In analogy to the application for electromagnetic waves, here we introduce the mechanism of destructive interference for elastic wave attenuation.

---

*Jiawen Xu is the corresponding author. (e-mail: jiawen.xu@seu.edu.cn).
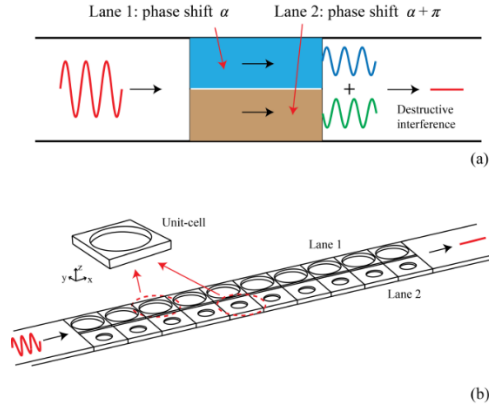
Fig.1. (a) Conceptual illustration of wave attenuation mechanism and (b) proposed PC system and the unit-cell.

The underlying physics is that the incident elastic wave is split and passes through the lanes in the plate with respective phase shifts (Figure 1(a)). The phase shifts of the waves in the adjacent lanes can be intentionally designed to meet the condition for destructive interference, resulting in the wave self-cancellation for vibration attenuation. For the sake of demonstration, we start from the equations of elastic waves in two adjacent lanes. Bloch theorem is adopted in the modeling. We characterize the elastic wave propagating within the two lanes in the following way:

$$u_1(x, k_1, t) + u_2(x, k_2, t) =$$
$$\tilde{u}_1(x, k_1)e^{i(k_1 \cdot x - \omega t + \varphi_1)} + \tilde{u}_2(x, k_2)e^{i(k_2 \cdot x - \omega t + \varphi_2)} \quad (1)$$

where $\mathbf{u}_1$ and $\mathbf{u}_2$ denote the displacement fields of the elastic waves in two adjacent lanes, respectively. $\tilde{\mathbf{u}}_1$ and $\tilde{\mathbf{u}}_2$ stands for the Bloch displacement function with the periodicity of the unit-cells, $\mathbf{x}$, $\mathbf{k}_1$, $\mathbf{k}_2$, φ_1, φ_2, ω and t denote position vector, wavevectors, phases, operating frequency and time. It is worth noticing that the incident waves in the two adjacent lanes have the same operating frequency and the amount of mechanical energy, i.e., they have very close amplitudes of displacements and wavenumbers. At the location $\mathbf{x}_0$ where the elastic wave converges, Equation (1) becomes

$$\left(u_1(x, k_1, t) + u_2(x, k_2, t)\right)\big|_{x=x_0}$$
$$= (\tilde{u}_1(x_0, k_1)e^{i\varphi_1} + \tilde{u}_2(x_0, k_2)e^{i\varphi_2})e^{i(k_2 \cdot x_0 - \omega t)} \quad (2)$$

Consider the fact that elastic waves in the two lanes have the same operating frequency, similar amplitudes and wavenumbers. Equation (2) yields

$$(\tilde{u}_1(x_0, k_1)e^{i\varphi_1} + \tilde{u}_2(x_0, k_2)e^{i\varphi_2})e^{i(k_2 \cdot x_0 - \omega t)}$$
$$\approx \tilde{u}(x_0, k)(e^{i\varphi_1} + e^{i\varphi_2})e^{i(k \cdot x_0 - \omega t)} \quad (3)$$

It can be obtained from Equation (3) that the amplitude of the elastic wave passing through the PC region is a function of the difference of their phases. Notably, the value of Equation (3) would be minimized when $|\varphi_1 - \varphi_2|$ approximates π. In other words, the elastic wave would be self-cancelled when the waves in the adjacent lanes have a phase difference of π, as conceptually illustrated in Figure 1(a). Therefore, the elastic wave has the potential to be attenuated by dispersion relation modulation in the PC lanes.

Without loss of generality, we assemble the proposed PC plate using coupled lanes consisting of staired unit-cells (Figure 1(b)). The staired unit-cell is a stair-shaped beam synthesized by two segments. Identical staired unit-cells are

connected in series to form one coupled PC lane. A small gap is introduced between the PC lanes to eliminate interactions between the lanes. Moreover, we coupled two adjacent lanes together and assemble many pairs of coupled lanes in parallel to form the PC plate. Each lane is formed by unit-cells with different dimensions, i.e., different dispersion relations and phase shift features. The difference of phase shifts would then produce the wave attenuation effects.

III. DISPERSION RELATION AND PHASE MODULATION

We start from analyzing the dispersion relation of a unit-cell. Finite element tool has been widely used in the investigations of mechanical phononic crystals. In the following finite element analysis using COMSOL 5.4, the unit-cell of the beam is formed by an aluminum substrate ($25 \times 25 \times 3$ mm^3) and one cylinder hole at the center of the surface of the unit-cell with a depth of 2 mm. The material mass density and Young's modulus are 2730 kg/m^3 and 62 GPa, respectively. We explore the dispersion relation in the first Brillouin-zone by sweeping of the wave number k from 0 to π/l_b, where l_b denotes the unit-cell' length, kl_b stands for the dimensionless wavenumber per unit-cell length.
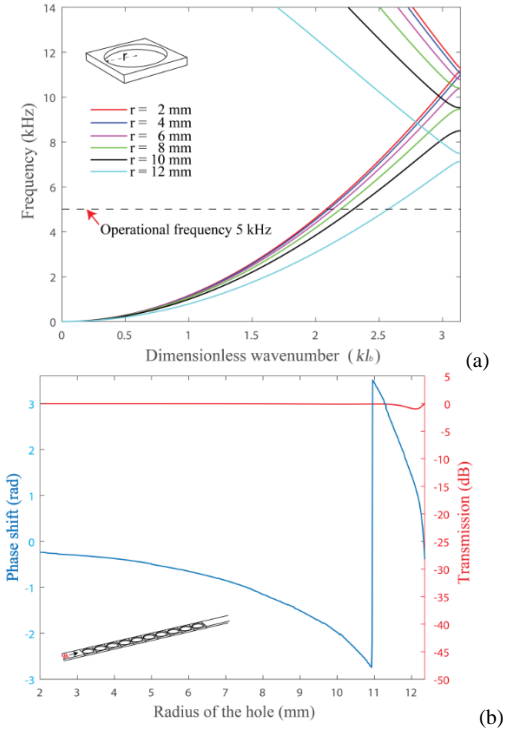
Fig.2. (a) Dispersion relations of the unit-cell with different r ; (b) Phase shift and transmission of a single phononic crystal lane at 5 kHz.

The unit-cell's dispersion curves with different radiuses (r) of the cylinder hole are shown in Fig. 2a. The dispersion relation of the phononic crystal shifts with the change of r. In other words, the incident elastic wave at a selected frequency is shortened or stretched given different holes. The dispersion relations implies the possibility of phase shift modulation for the elastic wave propagating through the medium and thus to achieve destructive interference in the design of phononic crystal beam by integrating various holes. Besides, Bragg scattering bandgaps of the phononic crystal systems with various holes can be observed in Fig. 2a.

Ten identical unit-cells are then serially connected to form a single functional lane of the finitely long phononic

crystal beam to understand the phase shift modulation by varying r (Fig. 2b). An operational frequency of 5 kHz below the Bragg scattering bandgaps is arbitrarily selected for demonstration. Boundary of perfectly matched layers (PMLs) are employed for the both ends of the beam. A transverse wave is incident from the left end of the beam. The characteristics of phase shift and the transmission of a single functional phononic crystal lane is analyzed at the chosen frequency by varying the value of the radius of the hole r. The phase shift plotted in the range of 2π and transmission diagrams are obtained by analyzing the responses in the far field as shown in Fig. 2b. It can be seen that the functional phononic crystal lane has a cumulative phase shift covering the full range of 2π. This indicates that the configured functional lane meets the requirement for rendering the destructive interference perfectly. Moreover, large transmission ratio above -2.5 dB with varying r at the operational frequency is noted.

IV. ELASTIC WAVE ATTENUATION CHARACTERISTICS
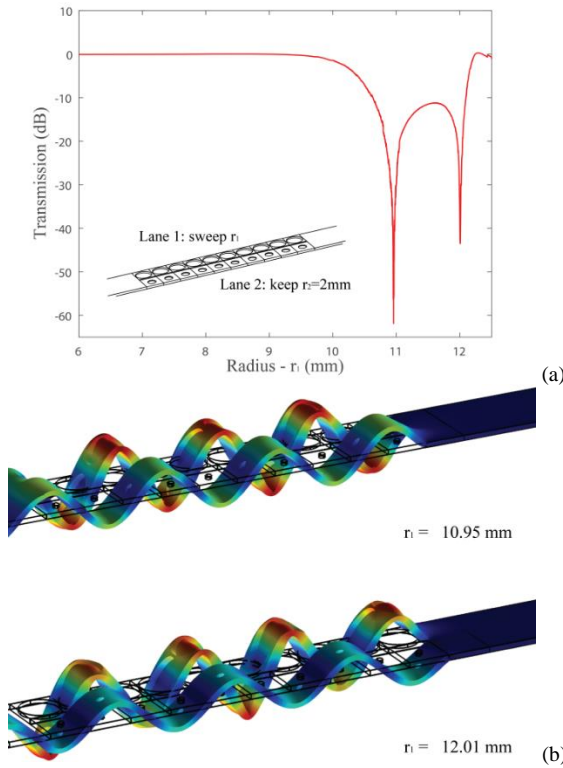
Fig.3. (a) Transmission diagram of dual lane phononic crystal beam by varying r_1 at 5 kHz; (b) Spatial responses for $r_1 = 10.95$ mm and $r_1 = 12.01$ mm.

We proceed to the synthesis of phononic crystal beam with two above-mentioned functional lanes. In the following analysis, r_1 and r_2 denote the radiuses of the holes for unit-cells in lane 1 and lane 2, respectively. We use such dual-lane phononic crystal beam for the demonstration of the wave attenuation mechanism resulting from the phase difference of the two lanes. Previous literatures proposed various designs for wave attenuation based on the mechanisms of elastic wave reflection or absorption. To exclude the wave attenuation effects owing to reflection or absorption, the lane 2 incorporates cylinder holes with $r_2 = 2$ mm, i.e., with a transmission of 1 (0dB) at 5 kHz (Fig. 2b).

The transmission diagram of the dual-lane phononic crystal beam at 5 kHz with varying r_1 is shown in Fig. 3a. It can be observed that promising wave attenuation is obtained at multiple values of r1, e.g., $r_1 = 10.95$ mm and $r_1 = 12.01$ mm. At these two values, the elastic waves transmitted through the two lanes have a phase difference of π, as shown in Fig. 3b. The phase difference yields the destructive interference and the consequent wave attenuation. It is worth mentioning that, since the elastic wave propagates in lane 1 has transmission above -2.5 dB (Fig. 2b), the wave attenuation here is indeed induced by the mechanism of destructive interference.

V. CONCLUSIONS

In summary, a phononic crystal beam based on the mechanism of destructive interference for broadband wave attenuation is demonstrated. The dual-lane phononic crystal beam consisting of cylinder holes integrated unit-cells is designed for realizing the proposed concept. Different phase shifts in each of the two lanes of the beam can be accomplished through proper selection of the radiuses of the cylinder holes. Numerical analysis confirms the significant wave attenuation owing to the destructive interference stemming from the different phase shift in the two lanes. The results confirm the extraordinary wave attenuation capacity of the proposed system.

REFERENCES

[1] Liu, Z., Zhang, X., Mao, Y., Zhu, Y. Y., Yang, Z., Chan, C. T., and Sheng, P., 2000, "Locally resonant sonic materials," *Science*, V289, 1734-1736.

[2] Fang, N., Lee, H., Sun, C., and Zhang, X., 2005, "Sub-diffraction-limited optical imaging with a silver superlens," *Science*, V308, 534.

[3] Landy, N. I., Sajuyigbe, S., Mock, J. J., Smith, D. R., and Padilla, W. J., 2008, "Perfect metamaterial absorber," *Physical Review Letters*, V100, 207402.

[4] Yang, J., Huang, M., Yang, C., Peng, J., and Chang, J., 2010, "An external acoustic cloak with N-sided regular polygonal cross section based on complementary medium," *Computational Materials Science*, V49, 9-14.

[5] Yoo, Y. J., Zheng, H. Y., Kim, Y. J., Rhee, J. Y., Kang, J. H., Kim, K. W,, Cheong, H., Kim, Y. H., and Lee, Y. P., 2014, "Flexible and elastic metamaterial absorber for low frequency, based on small-size unit cell," *Applied Physics Letters*, V105, 041902.

[6] Pendry, J. B., 2000, "Negative refraction makes a perfect lens," *Physical Review Letters*, V85, 3966.

[7] Climente, A., Torrent, D., and Sanchez-Dehesa, J., 2010, "Sound focusing by gradient index sonic lenses," *Applied Physics Letters*, V97, 104103.

[8] Zhao, J., Marchal, R., Bonello, B., and Boyko, O., 2012, "Efficient focalization of antisymmetric lamb wave in gradient-index phononic crystal plates," *Applied Physics Letters*, V101, 261905.

[9] Beck, B. S., Cunefare, K. A., Ruzzene, M., and Collet, M., 2011, "Experimental analysis of a cantilever beam with a shunted piezoelectric periodic array," *Journal of Intelligent Material Systems and Structures*, V22, 1177-1187.

[10] Xu, J., Li, S., and Tang, J., 2018, "Customized Shaping of Vibration Modes by Acoustic Metamaterial Synthesis," *Smart Materials and Structures*, V27, 045001.

[11] Yang, Z., Mei, J., Yang, M., Chan, N. H., and Sheng, P., 2008, "Membrane-type acoustic metamaterial with negative dynamic mass," *Physical Review Letters*, V101, 204301.

[12] Xu, J., Zhang, X., and Yan, R., 2020, "Coupled Piezoelectric Phononic Crystal for Adaptive Broadband Wave Attenuation by Destructive Interference", *Journal of Applied Mechanics*, 87(9).

[13] Kushwaha, M. S., Halevi, P., Dobrzynski, L., and Djafari-Rouhani, B., 1993, "Acoustic band structure of periodic elastic composites," *Physical Review Letters*, V71, 2022.

[14] Xu, J., Hu, G., Tang, L., Zhang, Y., and Yan, R., 2020, "Modeling and Analysis of Phononic Crystal with Coupled Lanes for Enhanced Elastic Wave Attenuation," *Journal of Vibration and Acoustic*, accepted.

[15] Tan, K. T., Huang, H. H., and Sun, C. T., 2014, "Blast-wave impact mitigation using negative effective mass density concept of elastic metamaterials," *International Journal of Impact Engineering*, V64, 20-29.

[16] Xu, J., and Tang, J., 2017, "Tunable Prism Based on Piezoelectric Metamaterial for Acoustic Beam Steering," *Applied Physics Letters*, V110, 181902.

[17] Chen, Y., Huang, G., and Sun, C. T., 2014, "Band Gap Control in an Active Elastic Metamaterial With Negative Capacitance Piezoelectric Shunting," *Journal of Vibration and Acoustics*, V136, 061008.

[18] Zhu, R., Liu, X., Hu, G., Sun, C., and Huang, G., 2014, "A chiral elastic metamaterial beam for broadband vibration suppression," *Journal of Sound and Vibration*, V333(10), 2759-2773.

[19] Hu, G., Tang, L., Das, R., Gao, S., and Liu, H., 2017, "Acoustic metamaterials with coupled local resonators for broadband vibration suppression," AIP *Advances*, V7, 025211.

[20] Xu, J., Yan, R., and Tang, J., 2018, "Broadening Bandgap Width of Piezoelectric Metamaterial by Introducing Cavity," *Applied Sciences*, V8(9).

[21] Fang, X., Wen, J., Bonello, B., Yin, J., and Yu, D., 2017, "Ultra-low and ultra-broad-band nonlinear acoustic metamaterials," *Nature communications*, V8(1), 1288.

[22] Deymier, P. , 2013. Acoustic Metamaterials and Phononic Crystals. Springer, Berlin Heidelberg.

Research Progress of Intelligent Lightning Protection Based on Internet of Things Technology

Qibin Zhou*
School of Mechanical Engineering and Automation
Shanghai University
Shanghai, China
zhouqibin@shu.edu.cn

Feipeng Tang
School of Mechanical Engineering and Automation
Shanghai University
Shanghai, China
747970181@qq.com

Xin Huang
School of Mechanical Engineering and Automation
Shanghai University
Shanghai, China
498289305@qq.com

Huahui Chen
Shanghai
Shanghai Meteorogical Bureau
Shanghai, China
605284052@qq.com

Abstract—Lightning protection is an important industry related to people's safety and interests. The traditional lightning protection needs to be combined with the Internet of Things. As the core equipment of the lightning protection system, the intelligent surge protector can limit the over-voltage flowing into the electrical system and discharge surge current. It protects equipment in the electrical system. The intelligent surge protector should be developed for safer and faster intelligent lightning protection. According to the development status of the Internet of Things, the working principle of intelligent surge protector is introduced. The framework of regional lightning intelligent monitoring and early warning system is presented from the perspective of architecture level. The machine learning method combined with the Internet of things was used to predict the life of the visitors. Realize intelligent lightning protection.

Keywords—internet of things, intelligent surge protector, regional lightning intelligent monitoring and early warning system, leakage current, life prediction

I. INTRODUCTION

With the advent of the Internet of Things era, modern buildings and industrial manufacturing industries are gradually becoming intelligent. A large number of microelectronic networks, automation equipment and computers are put into use. They have high integration, low working voltage and low working current. But the ability to withstand overvoltage and overcurrent is poor. If these devices are subjected to operating overvoltage or even lightning strikes, they may cause data loss and equipment damage. Therefore, preventing the equipment from being attacked by overvoltage and ensuring the reliable operation of the equipment system becomes a key and important task.

Surge protection device (SPD) is the core device of the lightning protection system, which protects the equipment in the electrical system. The safe and stable operation of SPD is essential to the lightning protection of electrical systems, so real-time detection of SPD is required to ensure the reliable operation of SPD. However, manual inspection requires a professional inspection team and a large number of inspection equipment, and the safe operation of the lightning protection system within the cycle cannot be fully guaranteed. For example, a surge protector installed in a building cannot work normally due to circuit connection problems or self-deterioration failure during the cycle. When the lightning surge arrives, the lightning current cannot be discharged into the ground in time, causing a short circuit failure.

In recent years, on the basis of the rapid development of sensors and Internet of things, the periodical maintenance mode of lightning protection system is gradually upgraded to the direction of 24-hour real-time monitoring of online monitoring system. There have been some researches on intelligent lightning protection based on Internet of things technology. The intelligent surge protector with lightning counting function is studied. It consists of surge protection module, lightning stroke sensor and lightning counting module. The lightning counting module and surge protection module are integrated into a whole, which solves the problems of large volume, low safety and high cost caused by the two sets of modules. An intelligent SPD for real-time monitoring of its own operation and performance is studied[1]. It can display the number of lightning strikes, leakage current, ambient temperature of SPD and lightning current waveform in real time. It can give an alarm when SPD has unexpected conditions. The working principle of intelligent SPD for collecting monitoring data is studied[2].The sampling, caching, reading and other aspects of real-time data in the actual working process are elaborated in detail.

Sec II describes the characteristics of two smart SPDS. Sec III introduces the structure of regional lightning intelligent monitoring and early warning system. Sec IV describes the use of machine learning to predict the residual life of SPD.

II. STRUCTURE AND OPERATION MECHANISM OF INTELLIGENT SPD

Intelligent lightning protection is an important part of the industrial Internet of things. SPD is one of the core equipment of low-voltage lightning protection system. It can be installed in low-voltage electrical circuits to protect the electrical system from continuous and safe operation under lightning strikes. It can also be installed at the front end of electronic equipment such as computers, communication and automatic control to shunt and inhibit overvoltage and protect equipment from damage [3]. It can also be installed between metal conductors for equipotential connection to balance the transient potential difference between different conductive systems caused by lightning electromagnetic pulse and avoid side lightning strike [4]. However, the failure of SPD installed in the system is often found in lightning protection engineering practice. The main reason is that qualified SPD products can be selected according to the performance characteristics of the protected equipment or system to protect the equipment within the expected service life of SPD in principle. However, due to the long-term work of SPD under power frequency voltage and continuous over-voltage

---

impact, leakage current gradually increases and temperature rise increases. With the increase of overvoltage times, the visitor's voltage changes and the protection effect become worse. What's more, the short-circuit failure of SPD leads to a sudden temperature rise and even a fire, as shown in Fig. 1. Therefore, the installation of SPD and reasonable wiring can no longer meet the requirements of modern lightning protection. Need to track and monitor the changes in SPD performance in real time [5]. Once SPD performance deteriorates or fails, the intelligent system should immediately send out an alarm to remind the staff to carry out maintenance or replacement so as to make the lightning protection system more stable. In recent years, the research on intelligent SPD has become a hot issue in the industrial Internet of things [6].

Fig.1. Photos of SPD failure and burn in fields

At present, many enterprises and research teams are engaged in the research and development of intelligent SPD. Although their structures are quite different, they can be divided into two categories based on monitoring principles.

A. Leakage Current Value Based on Monitoring SPD
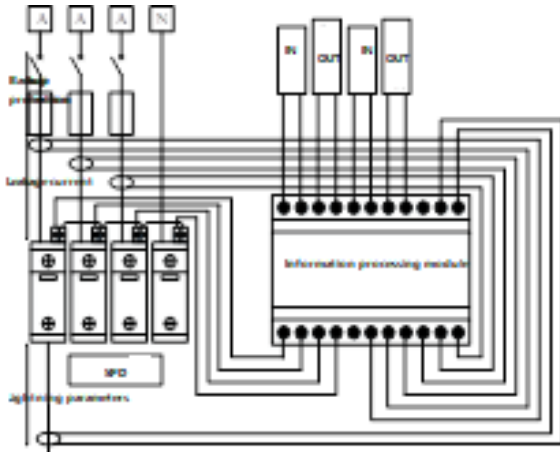
Fig.2. Structure diagram of seperated-type intelligent SPD

The intelligent SPD based on monitoring SPD leakage current value consists of SPD, leakage current monitoring sensor, lightning monitoring sensor, information processing module and backup protection, as shown in Fig. 2. The leakage current of intelligent SPD refers to the current passing through SPD under power frequency voltage[7]. When the SPD suffers from over-voltage shot and deterioration, the leakage current changes. When the leakage current is greater than a certain value, the heating of SPD increases rapidly. If the thermal protection device and backup protection device

are not cut off in time, a combustion event will be caused. Therefore, the leakage current can be used as an important parameter to judge whether SPD fails. The lightning monitoring sensor is responsible for surge counting and can record electrical parameters such as current waveform, charge and energy. The information processing module encodes the monitoring information collected by the sensor[8].

The visitors in SPD shows a high resistance state under low voltage. But there is still a leakage current of 100 microamperes under continuous power frequency voltage. The leakage current will increase as the SPD itself deteriorates. The temperature rise of SPD will also gradually increase [9]. To test the leakage current of a certain type of smart SPD, the steps are as follows: select 8 identical SPDs, and use 8/20μs, 20 kA impulse current to discharge each SPD 10 times at 20 °C. Then use the square wave impulse current of 1 to 5 kA to discharge the SPD of serial number 1 to 5 in a single impulse discharge. Under other conditions unchanged, the ambient temperature of serial numbers 6 and 7 was changed to 40°C and 60°C. Use a square wave impulse current of 5 kA to impulse the discharge. The leakage current of the SPD after the test is shown in Table I. It can be found that the leakage current increase of SPD is positively correlated with the overvoltage amplitude. The increase of SPD leakage current is positively correlated with the temperature of SPD operating environment.

TABLE I. TEST OF SPD LEAKAGE CURRENT

No.	surge current /kA	leakage current /mA
1	0	0.255
2	1	0.258
3	2	0.273
4	3	0.388
5	4	0.453
6	5	0.441
7	5	0.788

At present, such SPDs have been applied in substations in Beijing, Yunnan and other places. But according to several years of operation experience, it is found that SPD life cannot be accurately predicted simply by leakage current. The leakage current collected by current transformer is a full current, which includes resistive current and inductive current. The resistive current is the root cause of SPD heating, gradual deterioration and even failure. However, the unseparated full current does not reflect the actual abnormality. Although there is a means to separate full current, the added equipment not only increases the cost, but also needs to consider the lightning protection problem of the added equipment, which will also affect the building construction planning. At the same time, the key of the intelligent SPD 24-hour real-time monitoring technology is how to separate the useful dynamically changing resistive current. The realization of this technology requires a variety of algorithms, a variety of parameter combination calculation and a variety of sensor equipment synergy.

B. Based on Monitoring Surge Energy Value Flowing Through SPD

Monitoring the running state of the SPD based on the energy value requires time domain integration of the voltage and current flowing through the SPD to obtain the sum of surge energy flowing through the SPD. The energy threshold is set in combination with the lightning protection level of the protection system. When the total energy flowing through the SPD exceeds the threshold, a replacement warning is issued.

The basis for selecting the energy threshold is to carry out a square wave impact test on SPD, obtain the relationship between the surge times and the continuous working time of SPD under different amplitude currents, calculate the energy sustained by SPD at the same time, and select the working threshold by integrating the working temperature, humidity, lightning activity frequency and other conditions. Fig. 3 is a schematic diagram showing the change of the duration t_r of the square wave and the current magnitude I_{max} of certain SPD under a certain number of impact tolerance times.
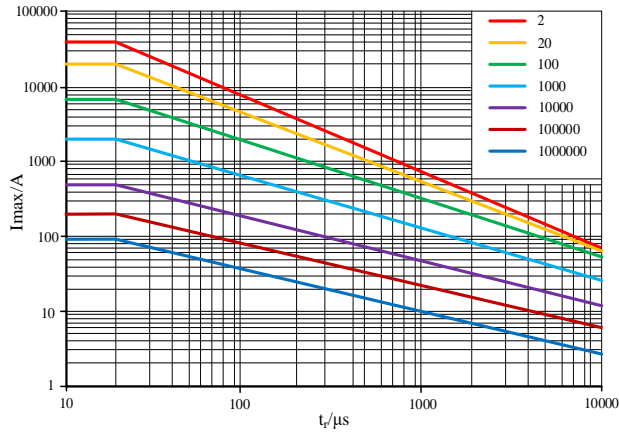
Fig.3. Relationship between duration t_r and current I_{max}

At present, the intelligent SPD selects the energy threshold according to the SPD energy tolerance and combined with the lightning protection level of the protected system, the regional average thunderstorm day, the environmental humidity and other factors. Adjust as external factors change. However, there are many uncertain factors, such as surge waveform, different amplitude and duration. In areas with fewer thunderstorm days, the SPD energy value is difficult to reach the threshold. However, the power frequency voltage applied persistently will inevitably lead to the deterioration of SPD. Smart SPD urgently needs more in-depth scientific research.

III. Intelligent Regional Lightning Monitoring and Warning System
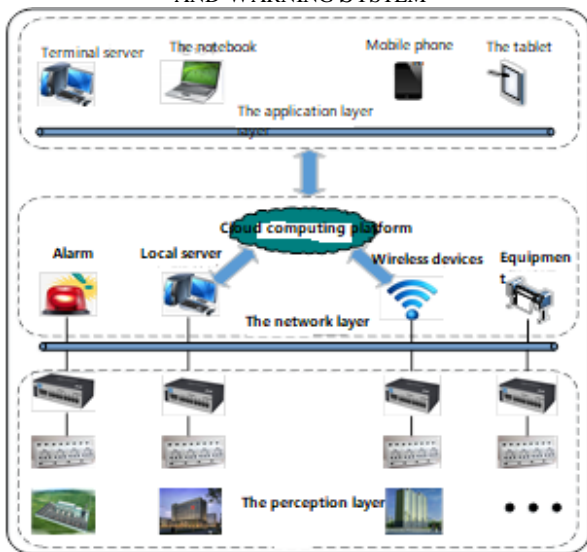
Fig.4. Intelligent regional lightning monitoring and warning system

The intelligent lightning protection system based on the Internet of things is called the lightning intelligent monitoring and early warning system. Each subsystem, such as substation, hospital and residential building, can be defined as a regional lightning intelligent monitoring and early warning system. The ultimate purpose of lightning protection is to protect people and things in the subsystem. Therefore, the regional lightning intelligent monitoring and early warning system is mainly analyzed. Its structure is shown in Fig.4.

The regional lightning intelligent monitoring and early warning system can be divided into perception layer, network layer and application layer on the architecture level. The three-tier architecture uses highly reliable and highly integrated equipment to interact with each other closely, enabling SPD to realize remote communication, telemetry and remote control functions to ensure the safe operation of substations, hospitals, residential areas, and high-precision equipment areas.

A. Perception Layer

The sensing layer [10] consists of SPD, sensor, backup protection, display screen, gateway, high-definition camera and other devices that collect external physical world data. The sensor measures voltage and current, senses external temperature, humidity, image and other information. Through field bus, Bluetooth, infrared and other short distance transmission technology, the information will be transmitted to the upper information integration terminal; USB interface can be installed to download or upload data to realize data transmission between monitoring station and SPD monitoring terminal. Meanwhile, administrator can manage SPD terminal through gateway. The display screen displays lightning strike times, overvoltage amplitude and life warning indication. Provide staff with the core parameters of the SPD, making it an intelligent terminal.

B. Network Layer

The network layer [11] solves the problem of data transmission obtained by the sensing layer. Network technologies such as long-distance wired and wireless communication technologies are applied to carry out data transmission through networks such as mobile communication networks, the Internet, enterprise intranets, various special line networks, local area networks, metropolitan area networks and the like, so as to realize interconnection between the sensing layer and the application layer. The data such as local alarm situation, lightning strike, lightning waveform, leakage current value and carrying capacity value are uploaded to the cloud computing platform through high-speed network.

C. Application Layer

The application layer is responsible for data integration and data processing[12]. It is composed of database server, intelligent monitoring information management platform, engineer maintenance platform, Web publishing service platform, database system, operating system and alarm software to realize the global management of intelligent SPD. The data from the monitoring layer is analyzed and processed, and the data is transmitted to the substation or mobile equipment through the cloud computing platform for alarm and linkage control. The database supports WEB browsing and APP interconnection modes, and interaction with the system can be realized in APP forms. The human-computer interaction here has gone far beyond the concept of human-computer interaction. But refers to feedback between various devices connected to application programs. Monitoring of protection elements of lightning protection system will not be

limited by factors such as time and place. Staff can obtain real-time information of work area at anytime and anywhere, and immediately take countermeasures such as disconnecting circuit breakers and reminding nearby personnel to change.

IV.   SPD LIFE PREDICTION

Under the background of Internet of things, the reliable operation of SPD is very important. So it is necessary to predict the service life of SPD. Visitors is the most important component of SPD. The key point of the life prediction research of surge protector is the research on the deterioration process of visitors[13].

Residual life prediction of visitors can be divided into two types: residual life prediction based on mechanism model and residual life prediction based on data drive. Residual life prediction based on mechanism model is the most accurate. But with the development of production technology, the deterioration mechanism of visitors becomes more and more difficult. As a large amount of prior knowledge is required, it is seldom used in practice. In the data-driven residual life prediction, the machine learning method is used to predict the residual life without understanding the specific mechanism of deterioration. A large number of relevant data collected by sensors are used to characterize the life of visitors, which has the characteristics of high efficiency and directness and is widely used in practice.

Machine learning method is used to predict the residual life of visitors. If the relationship between a single characteristic parameter and the life is adopted, the contingency of the result will be increased and the prediction effect will be poor. Therefore, various visitors deterioration related parameters should be used to predict the service life of visitors in practical engineering to effectively avoid accidents caused by SPD damage and improve lightning protection of the Internet of Things.

Five degradation data of visitors voltage, leakage current, capacitance, resistance and nonlinear coefficient are obtained from reference [14]. The experimental data were normalized. The results are shown in Fig 5.

$$x = (X^2 - X_{min})/(X_{max} - X_{min}) \tag{1}$$

In (1), X_{min} is the minimum value and X_{max} is the maximum value.

Because the nonlinear coefficient shows a decreasing trend with the number of shots. Visitors voltage, leakage current, capacitance and resistance are selected as the input parameters of the model. The nonlinear coefficient is taken as the output parameter. The nonlinear coefficient represents the remaining life of the visitors.
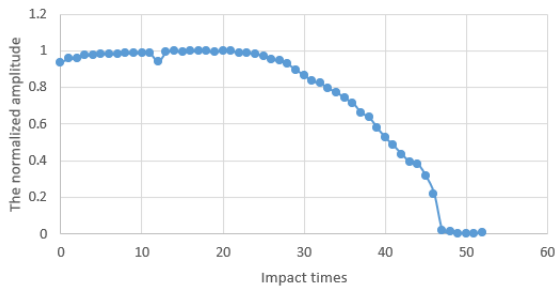
(a) Visitors voltage
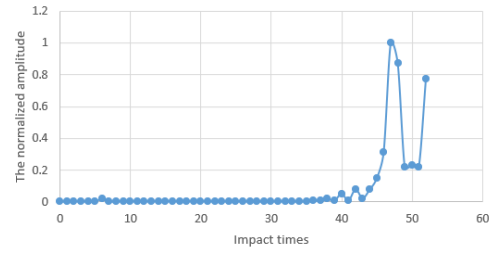
(b) Leakage flow
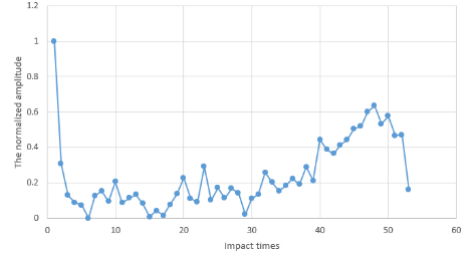
(c) Capacitance
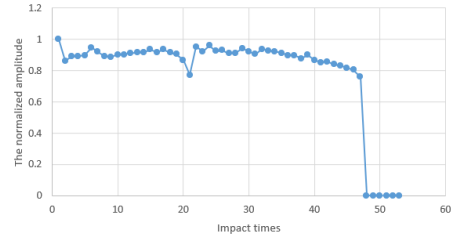
(d) Resistance
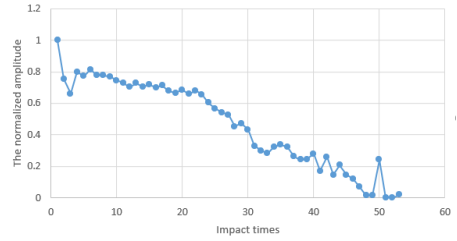
(e) Nonlinear coefficient

Fig.5. The normalized values of parameters varying with the number of shots

The nonlinear coefficient is fitted, as shown in Fig. 6
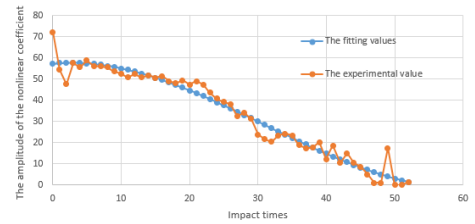
Fig.6. Nonlinear coefficient fitting diagram

$$y(i) = 6.48 \times 10^{-4} \times i^3 - 0.06 \times i^2 + 0.32 \times i + 57.3 \tag{2}$$

where i is the number of shot times.

The neural network model of the visitors is established. Since the output parameters are equal to four, there are four nodes at the input layer. The hidden layer nodes are selected as 7 according to (3). One output layer node.

$$L = \sqrt{m + n^2} + a \tag{3}$$

462

In (3), m represents the number of nodes in the input layer, n represents the number of nodes in the output layer. And, a is any parameter from 1 to 10.

Fig. 7 shows the comparison between the neural network prediction data and the experimental data. It can be seen that the neural network prediction of the remaining life of the varistor has a good performance and high accuracy.
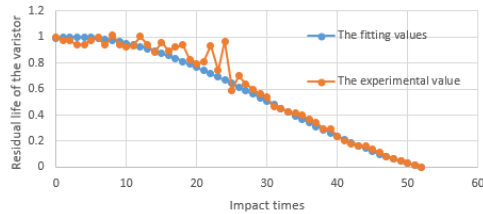
Fig.7. The normalized values of parameters varying with the number of shots

The machine learning method relies on a large number of experimental data to build a good model and has a strong dependence on data.

V. SUMMARY AND CONCLUSION

This paper combines Internet of things technology, big data technology and lightning protection technology. Research the development of intelligent lightning protection system under the Internet of Things technology. Lightning protection technology combined with Internet of things technology to realize online monitoring of SPD is a key step of intelligent lightning protection. This paper also combined with regional lightning warning system to form a relatively complete lightning protection system. The two SPD different monitoring principles can meet the needs of different scenes by selecting according to the actual scenes. The acquisition signal of lightning warning system is processed from the sensing layer to the application layer through the network layer to realize real-time online monitoring of SPD state and lightning stroke count. Life prediction of SPD is very critical to intelligent SPD and lightning protection technology. Life prediction of SPD through collected data and machine learning method can effectively predict the life of SPD. Providing reference basis for intelligent SPD and intelligent lightning protection, timely replacement of components,

reducing the damage caused by lightning stroke, and also reducing the waste of manpower and material resources.

REFERENCES

[1] Y. Xiucheng, H. Tingjie, and W. Bo, "A Brief Analysis of SPD Intelligent Regulatory Early Warning System based on Internet of Things technology," Building Electrical, pp.57-60, 2016.

[2] W. Yong, "Working Principle and Application of SPD Intelligent Monitoring Data Collector," Power System Communication, pp. 79-81, 2012.

[3] G. Liu, H. Shi, and R. Shi, "Analysis of the suppression effect of short-circuit line SPD in different frequency Bands on Lightning Wave," A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, pp.68–73, 1892.

[4] P. Colella, E. Pons, and R. Tommasini, "Dangerous touch voltages in buildings: The impact of extraneous conductive parts in risk mitigation," Electric Power Systems Research, pp.263-271, 2017.

[5] T. Liu, A. Shen, and J. Hu, "The application of collaborative business intelligence technology in the hospital SPD logistics management modle," Iranian Journal of Public Health, pp.744-754, 2017.

[6] Z. Huang, R. Wang, and X. Li, "Geometry-aware similarity learning on SPD manifolds for visual recognition," IEEE Transactions on Circuits & Systems for Video Technology, pp.1-3, 2016.

[7] Q. Lu, X. Li, and L. Xiong, "Method for on-line detection of leakage current of power supply SPD," IEEE, International Symposium on Electromagnetic Compatibility, pp.1-5, 2017.

[8] K. Park, G. Wang, and S. Hwang, "Deterioration characteristics and an on-line diagnostic equipment for surge protective devices," Journal of the Korean Institute of Electrical and Electronic Material Engineers, pp.635-640, 2016.

[9] J. Yangzhe, L. Chen, and Y. Sun, "Application on capacitance during the degradation of ZnO varistor," High Voltage Engineering, pp.2167-2172, 2010.

[10] F. Guo, X. Zeng, and W. Hui, "Research on temperature field distribution and thermal failure of MOV used in SPD based on the multi-physics coupling finite rlement method," Insulators & Surge Arresters, pp.33-36, 2017.

[11] B. M. Jordi and K. Piotr, "Conception of ID layer performance at the network level for internet of things," Personal & Ubiquitous Computing, pp.465-480, 2014.

[12] D. Han, J. Zhang, and Y. Zhang, "Convergence of sensor networks/internet of things and power grid information network at aggregation layer," International Conference on Power System Technology, pp.1-6, 2010.

[13] W. Yuanfang, Y. Xianglian and L. Chun, "A method for the lifetime prediction of MOV," 33rd Symposium on Electrical and Ele, pp.757-760, 2001.

[14] J. Shaoyun, "Study on residual life Model of Medium-voltage varistor of Railway Surge Protector," China Academy of Railway Sciences, 2019.

Parameter Study of the Variable Reluctance Energy Harvester for Smart Railway Axle Box Bearing

Yun Gong
The State Key Lab of Mechanical Transmissions
Chongqing University
Chongqing, China
gy5401@163.com

Sijia Wang
The State Key Lab of Mechanical Transmissions
Chongqing University
Chongqing, China
2750683017@qq.com

Zhengqiu Xie
College of Mechanical Engineering
Chongqing University of Technology
Chongqing, China
zhengqiuxie@163.com

Wenbin Huang*
The State Key Lab of Mechanical Transmissions
Chongqing University
Chongqing, China
whuang@cqu.edu.cn

Abstract—**Self-powering capability for bearing health monitoring systems is an appealing ideal, which has attracted considerable attention but difficult to realize. This paper proposes a variable reluctance energy harvester (VREH) integrated within in smart railway axle box bearing. The effects of key parameters are numerically discussed to obtain the optimal geometrical parameters of the VREH under the limited inner space of the bearing. The experimental results verify that the optimization design of the VREH is practicable and effective. Under the rotational speed of 1200 r/min and impedance matching, the RMS voltage and power of the optimal load reaches 1.54 V and 39.53 mW, respectively.**

Keywords—self-power, variable reluctance, health monitoring, theoretical modeling, finite element analysis, bearing

I. INTRODUCTION

Sensors and sensor systems play an increasingly important role in the operation, management and maintenance of industrial machinery [1]. Conventional batteries provide power for such devices is inevitable to result in regular recharging and replacement, and the pollution of environment [2]. Electromagnetic [3,4], electrostatic [5,6] and piezoelectric [7,8] energy harvesters have been designed to convert the rotating motion into energy and power the wireless sensor and sensor system embedded on such rotating parts as bearings.

In various energy harvesters, the electromagnetic energy harvester (EMEH) remains one of the leading topics because of the high energy density, easy fabrication and low cost [9]. Xu *et al.* [10,11] investigated a M-shaped VREH at low angular velocities. By optimizing the parameters affecting the performance of VREH through simulation and experimental analysis, the power density of VREH is demonstrated to be able to reach 157.29 mW/cm3. Kim *et al.* [12] combined a rotational module triggered by finger with an array of disc Halbach magnets to harvest electromagnetic energy. The proposed energy harvester could generate an open-circuit voltage of 1.39 V at an average power of 7.68 mW, with an optimal load of 36 Ω at an input frequency of 3 Hz. Kroener *et al.* [13] proposed a VREH for railroad monitoring applications. By conducting the finite element analysis (FEA) and experiments, a maximum energy output of 131 μJ per pulse was obtained under a train wheel passing speed of 81.5 km/h, corresponding to a mean output power of 5.9 mW. Zhang *et al.* [2,14] designed a EMEH with circular Halbach for bearing, and verified the effectiveness of proposed model

numerically and experimentally. Under the rotational speed from 600-1000 r/min, the voltage of 2.79-4.59 V and the maximum average power of 50.8-131.1 mW can be obtained by the enhanced harvester.

This paper proposed a theoretical modal of VREH that extracted energy from the railway axle box bearing. Different from previous studies which put the energy harvester in the bearing end cover as an external accessory, we integrated the VREH with the bearing. According to the limited interior space of smart railway axle box bearing, to enhance the performance of VREH, the influences of various structure parameters on the output response are investigated through numerical simulation and experiments. The remainder of this paper is organized as follow: section 2 proposes the design and model of VREH; section 3 depicts the numerical analysis of VREH and analyzes the effects of system parameters on output response by comparing the simulation and theory; the effectiveness of VREH is verified by experiments in section 4; finally some conclusions are drawn in section 5.

II. MODELING OF VREH

A. Structure of VREH
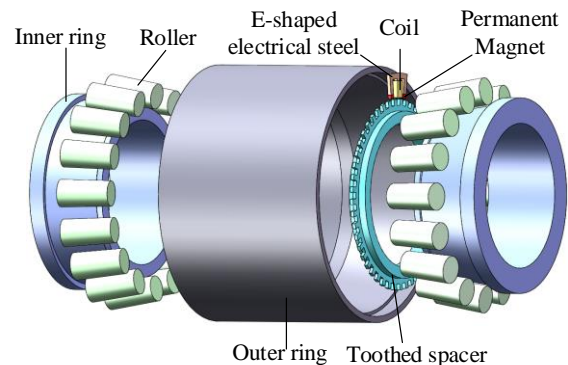
Fig. 1. Schematic of the VREH in bearing.

Fig. 1 illustrates the proposed VREH in the smart railway axle box bearing. It comprises an E-shaped electrical steel, two permanent magnets, a coil, a toothed spacer, two inner rings, rollers and an outer ring of bearing. The E-shaped electrical steel, permanent magnets, coil and toothed spacer constitute a VREH. The E-shaped electrical steel is embedded in the outer ring of the bearing, and there is an air gap between the E-shaped electrical steel and the toothed spacer. The two permanent magnets with the same polarity are respectively

located at the ends of two outer leg of the E-shaped electrical steel, and there is an air gap between the permanent magnets and teeth of the toothed spacer. The coil is wrapped around the middle leg of the E-shaped electrical steel.

(a) Aligned position   (b) Unaligned position
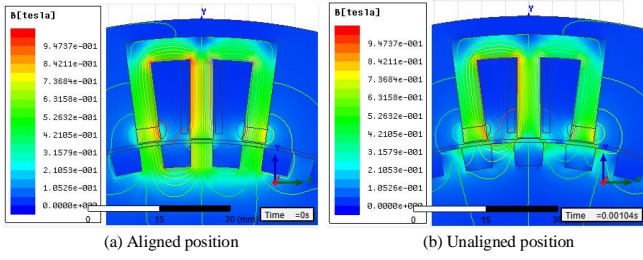
Fig. 2. Structure of VREH and the magnetic flux density distribution at its two extreme position.

As the smart railway axle box bearing rotates, there is relative motion between the toothed spacer and E-shaped electrical steel, which makes the magnetic flux through the coil change with time. Therefore, according to Faraday's law of electromagnetic induction, the induction electromotive force will be generated by the coil. Since E-shaped electrical steel and toothed spacer are made of materials with high permeability, there are two extreme cases of magnetic flux within the coil, as shown in the Fig. 2. At the aligned position, the teeth of the toothed spacer are aligned with the legs of E-shaped electrical steel. The magnetic field induced by the two permanent magnets passes through the coil along the E-shaped silicon steel, and then passes through the air gap with the teeth of the toothed spacer to form the minimum magnetic path resistance, as shown in Fig. 2a. At the unaligned position, the teeth of the toothed spacer are not aligned with the legs of E-shaped silicon steel, and the magnetic field induced by the two permanent magnets passes through the coil along the E-shaped silicon steel, and then passes through the air gap and the slot instead of the teeth of the toothed spacer, forming the maximum magnetic path resistance, as shown in Fig.2b.

B. Modeling of VREH

(a) Cross-section of smart railway axle box bearing   (c) Cubic magnet
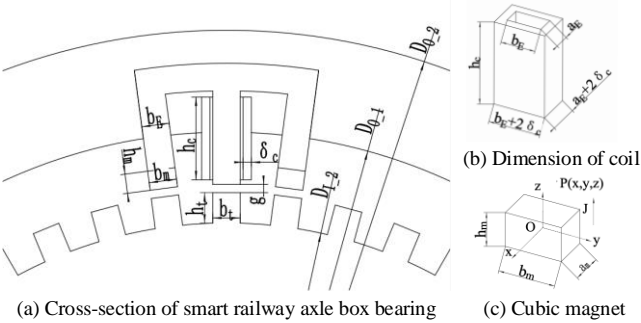
Fig. 3. Definitions of geometric parameters in VREH.

With the rotation of the smart railway axle box bearing, the internal VREH converts the kinetic energy into the electric energy. The cross-section of smart railway axle box bearing without showing the inner ring and the roller is plotted in Fig. 3a. According to Faraday's law of electromagnetic induction, the induced electromotive force of the coil can be expressed as

$$U(t) = -N_c \frac{d\phi}{dt} \tag{1}$$

where N_c is the turns of coil, Φ is the magnetic flux through the coil, $d\Phi/dt$ is the gradient of magnetic flux.

As shown in Fig. 3b, according to the geometry of the coil, the number of turns of the coil can be expressed as

$$N_c = \frac{\chi \delta_c h_c}{\frac{\pi}{4} d_c^2} \tag{2}$$

where χ, δ_c, h_c and d_c are the filling rate, width, height and the wire diameter of the coil, respectively.

The frequency of the output voltage of the coil is linear with the speed of the bearing and the number of teeth of the toothed spacer, which can be expressed as

$$f_c = \frac{n}{60} N_t = \frac{n}{60} \frac{\pi(D_{L\_2} + h_t)}{2b_t} \tag{3}$$

where n (unit: r/min) is the speed of the bearing, N_t is the number of teeth, D_{L_2} is the outer diameter of the bearing inner ring, h_t and b_t are the height and width of the tooth of the toothed spacer, respectively.

As the time-varying flux in (1) is a uniform sinusoidal voltage, the peak-peak value of the output voltage V_{pp} generated by the coil can be expressed as

$$V_{pp} = -2N_c f_c \Delta\phi = -N_c \frac{\phi_{max} - \phi_{min}}{\frac{T_c}{2}}$$
$$= -2N_c f_c \xi_m (\frac{1}{\tau_{min}} - \frac{1}{\tau_{max}}) \tag{4}$$

where $\Delta\Phi$ is the difference between the maximum flux Φ_{max} and minimum flux Φ_{min} of the coil, T_c is the period of the induced electromotive force of the coil, $\zeta m = H \cdot h_m$ is the magnetomotive force generated by the permanent magnet, H is the magnetic field strength of the magnet, h_m is the height of the magnet, and τ_{min} and τ_{max} are the minimum and maximum reluctance, respectively.

In the magnetic path of the proposed VREH, the permeability of the material used in the E-shaped electrical steel and the toothed spacer are relatively larger. When calculating the reluctance, only the magnets and air gap between the permanent magnet and the toothed spacer is considered, which can be expressed as

$$\tau = \frac{g}{\mu_0 \mu_g a_t b_t} + \frac{h_m}{\mu_0 \mu_m a_m b_m} \tag{5}$$

where $g=g_{min}$ is the air gap at the maximum flux of the coil, the air gap is $g=g_{min}+h_t$ at the minimum flux, and μ_g is the relative permeability of the air, a_t and b_t are the thickness and width of any tooth of the toothed spacer, μ_m is the relative permeability of the permanent magnet, and a_m and b_m are the thickness and width of the permanent magnet, respectively.

Since the magnetization direction of the permanent magnet is along the radial direction of the bearing, a coordinate system is established with the original point located at the center of the permanent magnet, as shown in Fig. 3c, according to the magnetic load theory [15] and the magnetic node theory [16], the magnetic field strength of given point $P(x,y,z)$ along the magnetization direction, which produced by this magnet can be expressed as

$$H = \frac{J}{4\pi\mu_0} \sum_{k=0}^{1} \sum_{q=0}^{1} \sum_{m=0}^{1} tg^{-1}(\frac{U_q V_m}{W_k r}) \tag{6}$$

where, J is the uniform magnetization along the magnetization direction, μ_0 is the vacuum permeability, and $U_q = x - (-1)^q \cdot a_m$, $V_m = y - (-1)^m \cdot b_m$, $W_k = z - (-1)^k \cdot h_m$, $r = \sqrt{U_q^2 + V_m^2 + W_k^2}$.

Substituting (3), (5), and (6) into (4), the expression of the peak-peak voltage V_{pp} generated by the coil is easily obtained as

$$V_{pp} = -\frac{\chi \delta_c h_c}{\pi d_c^2} \frac{n}{60} \frac{(D_0 + h_t)}{b} \frac{J}{\mu_0} \sum_{q=0}^{1} \sum_{m=0}^{1} \sum_{k=0}^{1} tg^{-1}\left(\frac{U_q V_m}{W_k r}\right)$$
$$\cdot \frac{h_m h_t}{(\mu_M l_g + \mu_g h_m) + (\mu_m(l_g + h_t) + \mu_g h_m)} \quad (7)$$

where $b = b_t = b_m$.
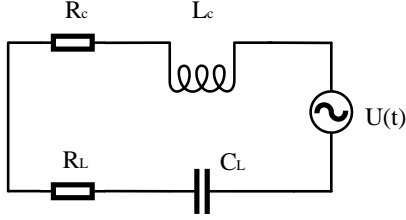
Fig. 4. External circuit with impedance matching.

Due to the high speed of railway bearing, the inductance of the coil cannot be neglected. In order to obtain the optimal output power from the VREH, the coil is equivalent to a series of inductance and resistance, after impedance matching, as shown in Fig. 4, the RMS voltage of the load resistance $U_{L,rms}$ can be expressed as

$$U_{L,rms} = \frac{R_L}{Z_{total}} \frac{\sqrt{2}}{2} V_{pp} = \frac{\sqrt{2}}{4} V_{pp} \quad (8)$$

where, $Z_{total} = R_L - j/2\pi \cdot f_c \cdot C_L + R_c + j \cdot 2\pi \cdot f_c \cdot L_c$ is the total impedance, $C_L = 1/(2\pi \cdot f_c)^2 \cdot L_c$ is the matched capacitance, and $R_L = R_c$ is the matched resistance.

Substituting (7) into (8), the expression of the effective value of voltage of the load resistance $U_{L,rms}$ can be easily obtained as

$$U_{L,rms} = -\frac{\sqrt{2}}{4} \frac{\chi \delta_c h_c}{\pi d_c^2} \frac{n}{60} \frac{(D_0 + h_t)}{b} \frac{J}{\mu_0} \sum_{k=0}^{1} \sum_{q=0}^{1} \sum_{m=0}^{1} tg^{-1}\left(\frac{U_q V_m}{W_k r}\right)$$
$$\frac{h_m h_t}{(\mu_M l_g + \mu_g h_m) + (\mu_m(l_g + h_t) + \mu_g h_m)} \quad (9)$$

where the internal resistance of the coil can be expressed as

$$R_c = 4\rho_c \frac{N_c \overline{C}}{\pi d_c^2} = 4\rho_c \frac{N_c(a_t + b_t + 2\delta_c)}{\pi d_c^2} \quad (10)$$

Under impedance matching, the RMS power $P_{L,rms}$ of the load can be expressed as

$$P_{L,rms} = \frac{U_{L,rms}^2}{R_c} \quad (11)$$

According to (9) and (11), the RMS voltage of the load is linear with the speed of the bearing n, the width δ_c and height h_c of the coil, while RMS power is quadratic with those parameters. Moreover, the value of RMS voltage and power of the load are related with the minimum air gap g_{min}, the height of the permanent magnet h_m, the height of the toothed spacer h_t and the tooth width of the permanent magnet and the toothed spacer b. Due to the limited inner space of the bearing, the summation of coil height h_c and the tooth height h_t of the toothed spacer are set as a fixed value, namely $h_c + h_t = C$, and

the thicknesses of E-shaped electrical steel, magnets, toothed spacer are the same, 3mm, and other parameters are listed in the Table I according to the actual working condition of the VREH. Then the RMS voltage and power of the load as functions of the coil height h_c and the tooth height h_t of the toothed spacer are shown in the Fig. 5a and Fig. 5b, respectively.

(a) RMS voltage   (b) RMS power
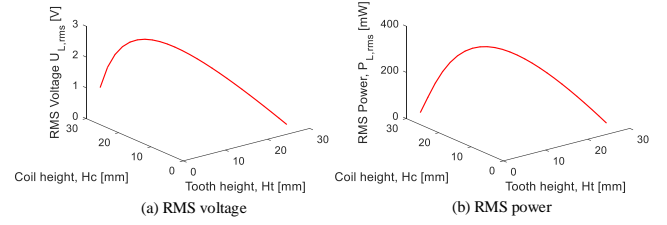
Fig. 5. Theoretical calculation of the response of load with $h_c + h_t = 27$mm.

Fig. 5 shows that the optimal value of RMS power of the load can be obtained with $h_c = 17$ mm and $h_t = 10$ mm.

TABLE I. GEOMETRICAL PARAMETERS AND MATERIAL SPECIFICATION

Domain	Parameters	Value
Magnet	h_m	2-10 mm
	b_m	5-15 mm
	Material	N33AH
Coil	h_c	1-27 mm
	d_c	0.15 mm
	δ_c	1.5 mm
	χ	0.45
E-shaped electric steel	b_E	5-15 mm
	Material	B35A230
Toothed spacer	h_t	1-27 mm
	b_t	5-15 mm
	Material	Steel 45
Air	g_{min}	1-4 mm
	μ_g	1
Inner ring and outer ring of bearing	D_{I_2}	150 mm
	D_{O_1}	193 mm
	D_{O_2}	230 mm
	Material	GCr15

III. NUMERICAL ANALYSIS OF VREH

A. Numerical Analysis Setup

To verify the validity and accuracy of the theoretical model, the finite element method is used to compare the simulation results with the theoretical calculation. Therefore, considering the influence of bearing outer ring and the symmetry of smart railway axle box bearing, the VREH structure is numerically analyzed based on 2D electromagnetic model. All the simulations are implemented in the Maxwell Ansoft. The parameters used in the simulation are consistent with those used in the previous theoretical calculation.

Due to the high speed of the bearing, the inductance of the coil cannot be neglected. In the simulation process, the external circuit of the coil is shown in Fig. 4. Where, R_c and L_c are the internal resistance and inductance of the coil and R_L and C_L are the matched resistance and inductance of the load, respectively.

B. Effect of Key Parameters

According to the numerical results and the inner space limitation of the bearing, the influence of the air gap g_{min}, permanent magnet height h_m, and tooth width of the permanent magnet and toothed spacer b are investigated under the rotating speed of 600r/min.
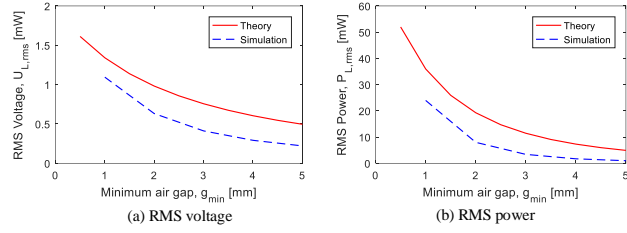
Fig. 6.   The RMS response of the load under different minimum air gap.

The relationship between the RMS voltage and power of the load with the minimum air gap are shown in Fig. 6a and Fig. 6b, respectively. The trend of simulation results is consistent with that of theoretical calculation, which both decrease rapidly with the increase of the air gap.
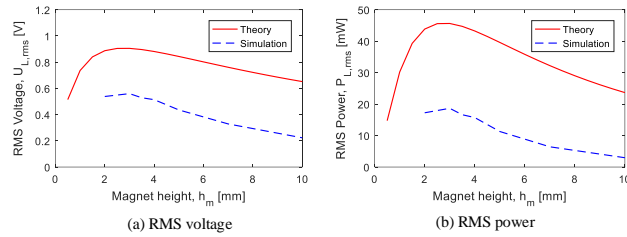
Fig. 7.   The RMS response of the load under different permanent magnet height.

Fig. 7a describes the RMS value of voltage and power of the load varying with the height of permanent magnet. The trend of simulation results is consistent with that of theoretical calculation, which increases first and then decreases with the increase of permanent magnet height. The optimal value occurs at h_m=3 mm is with a low rate of change.
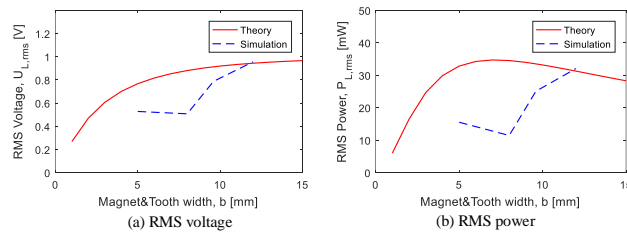
Fig. 8.   The RMS response of the load under different the width of the permanent magnet and the teeth of toothed spacer.

Fig. 8 represents the RMS voltage and power of the load with different width of the permanent magnet and the teeth of toothed spacer. The theoretical results increase with the increase of the width of permanent magnet and teeth of the toothed spacer, while the simulation results decrease first and then increase. This difference can be caused by errors in the

magnetic field distribution under the two-dimensional simulation model.

Simulation and theoretical results shows that the trends with the key parameters are the same, but the simulation results are smaller than the theoretical calculation, the difference is close to 20%. This is because we assume that the magnetic path of the VREH passes through the bottom of the slot of the toothed spacer at unaligned position. In fact, much of the magnetic flux leaks to the edges of the teeth, thus most of the magnetic flux cannot passes the slot as assumed, which can be verified in Fig 2b. Therefore, the modeled magnetic resistance difference between the aligned and unaligned positions is higher than the real case, resulting into a higher output compared with the simulation.

IV.   EXPERIMENT VERIFIVICATIONS

A. Experimental Setup
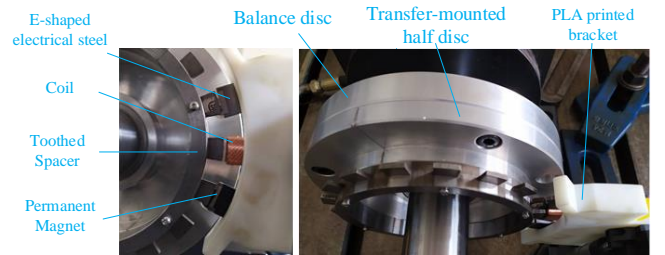
Fig. 9.   The prototype and installation of VREH.

The VREH prototype is shown in Fig. 9, which consists of a toothed rotor, an E-shaped electrical steel, two permanent magnets, a coil and a 3D printed bracket. The permanent magnets and coil are bonded to the legs of E-shaped electrical steel, and the E-shaped electrical steel is supported by the PLA printed bracket.
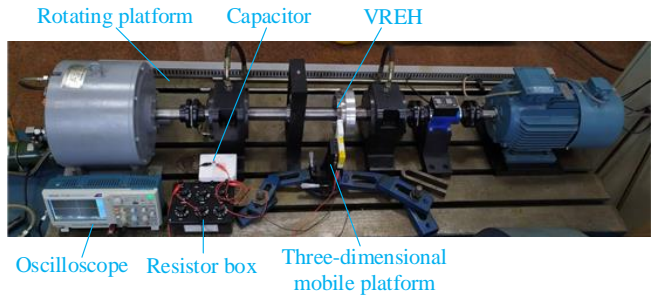
Fig. 10. Experimental platform of VREH.

TABLE II.   OPTIMAL GEOMETRICAL PARAMETERS

Domain	Parameters	Value
Magnet	h_m	3 mm
	b_m	14.6 mm
Coil	h_c	17 mm
Toothed spacer	h_t	10 mm
	b_t	14.6 mm
Air	g_{min}	1 mm

To verify the practicability and effectiveness of the proposed VREH and simulate the motion of VREH in the bearing, the VREH prototype is installed on a rotating platform, as shown in Fig. 10. The bracket is screw-mounted on the three-dimensional mobile platform, so the position of

E-shaped electrical steel can be adjusted, to align its legs with the toothed spacer with a desired air gap. In order to avoid the eddy current effect, E-shaped electrical steel is processed by wire cutting, then stacked and welded by 6-layers of electrical steel laminations, with an overall thickness of 3.15mm. The toothed rotor is also screw-mounted on the balance disc of the platform. The toothed rotor is actuated by the rotating platform with shaft and the rotating speed can be adjusted. According to the previous parameter optimization study, structural parameters listed in Table II are used for the prototype.

To obtain the optimal output power of the load, the output end of the coil is connected in series with a matching capacitor and resistor box, and the internal resistance and inductance of the coil was measured by a LCR-meter. All the experimental data are collected by the oscilloscope.

B. Voltage Response
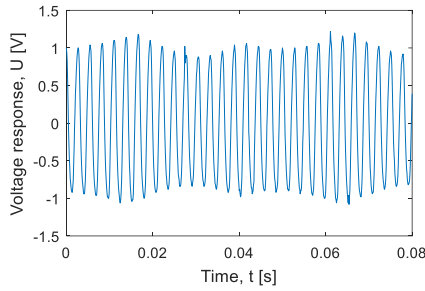
Fig. 11. Voltage response of the load under 600 r/min.

The minimum air gap is selected to 1 mm, and a 50 Ω load is connected at the output end of the coil in series. At the speed of 600 r/min, the voltage response of the load is shown in Fig. 11. As shown in the figure, the maximum value of the voltage response fluctuates periodically, and it is due to the construction of the experimental platform. There exist an eccentric errors between the toothed rotor and the shaft, causing a gap difference of 0.44mm between the maximum and minimum air gaps.
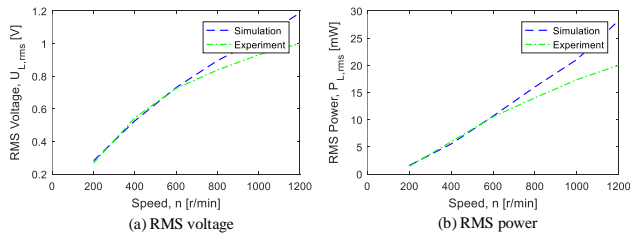
Fig. 12. The RMS response of the load under different rotational speed.

Figure 12 illustrates a linear relationship between the RMS voltage and rotation speed, which is consistent with the simulation results. Due to the existence of the air gap difference, the experimental results are slightly smaller than the simulation at higher speed.

C. Response Range

In order to obtain response range of the load, the values of RMS voltage and power of the load are obtained under the speed of 1200 r/min. Fig. 13 compares the RMS voltage and power of the load with and without matched capacitance. Without the matching capacitor, the RMS voltage of 1.42 V and the RMS power of 22.4 mW are obtained, with an optimal load of 90 Ω. With the matching capacitor, the RMS voltage

of 1.54 V and the RMS power of 39.53 mW can be achieved, with an optimal load of 60 Ω.
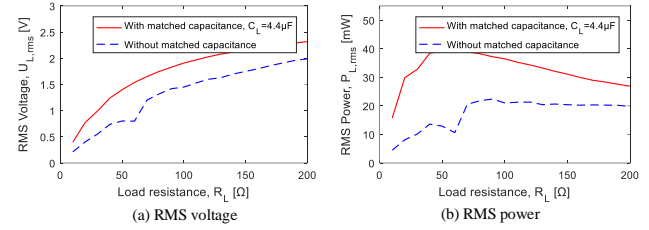
Fig. 13. The RMS response of the load with and without matched capacitor.

V. CONCLUSION

A VREH with enhanced configuration is proposed to support the wireless sensor in smart railway axle box bearing, which converts rotational motion into electric energy. The design device consists of an E-shaped electrical steel, two permanent magnets, a coil, a toothed spacer, two inner rings, rollers and an outer ring of bearing. Under the limitation of inner space of the bearing, the effect of several key parameters were investigated through theoretical modeling and finite element simulation, and optimal parameters of VREH were obtained. The experimental prototype was fabricated based on the optimal geometrical parameters, and rotating experimental platform is constructed to verify the accuracy of output prediction. In addition, the effect of the key system parameters and external excitation on output response characteristics are investigated. Experimental results indicate that the optimized VREH can generate a voltage of 1.54 V and a power response of 39.53 mW under the rotation speed of 1200 r/min.

REFERENCES

[1] D. Kwon, M. R. Hodkiewicz, J. Fan, T. Shibutani, M. G. Pecht, "IoT-based prognostics and systems health management for industrial applications," IEEE Access, vol. 4, pp. 3659–3670, 2016.

[2] Y. Zhang, J. Cao, W.-H. Liao, L. Zhao, J. Lin, "Theoretical modeling and experimental verification of circular Halbach electromagnetic energy harvesters for performance enhancement," Smart Materials and Structures, vol. 27, no. 9, pp. 095019, Aug. 2018.

[3] K. Fan, S. Liu, H. Liu, Y. Zhu, W. Wang, D. Zhang,"Scavenging energy from ultra-low frequency mechanical excitations through a bi-directional hybrid energy harvester," Applied Energy, vol. 216, pp: 0306-2619, 2018.

[4] Y. Zhang, T. Wang, A. Luo, Y. Hu, X. Li, F. Wang, "Micro electrostatic energy harvester with both broad bandwidth and high normalized power density," Applied Energy, vol. 212, pp: 0306-2619, 2018.

[5] B. Vysotskyi, D. Aubry, P. Gaucher, X. L. Roux, F. Parrain, E. Lefeuvre, "Nonlinear electrostatic energy harvester using compensational springs in gravity field," Journal of Micromechanics and Microengineering, vol. 28, no. 7, pp: 28-074004, 2018.

[6] A. Erturk, J. Hoffmann, D. J. Inman, "A piezomagnetoelastic structure for broadband vibration energy harvesting," Applied Physics, vol. 94, pp: 254102, 2009.

[7] I. Izadgoshasb, Y. Y. Lim, N. Lake, L. Tang, R. V. Padilla, T. Kashiwao, "Optimizing orientation of piezoelectric cantilever beam for harvesting energy from human walking," Energy Conversion and Management, vol. 161, pp:0196-8904, 2018.

[8] M. F. Daqaq. "On intentional introduction of stiffness nonlinearities for energy harvesting under white Gaussian excitations," Nonlinear Dynanics, vol. 69, pp: 1063–1079, 2012.

[9] L. Chen, T. Yuan, "Harmonic Balance in the Dynamic Analysis of Circular Composite Plate Harvester," Procedia IUTAM, vol. 22, pp: 2210-9838, 2017.

[10] Y. Xu, S. Bader,B. Oelmann, "Design, modeling and optimization of an m-shaped variable reluctance energy harvester for rotating applications," Energy Conversion and Management, vol. 195, no. February, pp: 1280–1294, 2019.

[11] Y. Xu, S. Bader, B. Oelmann, "A Survey on Variable Reluctance Energy Harvesters in Low-Speed Rotating Applications," in IEEE Sensors Journal, vol. 18, no. 8, pp. 3426-3435, 15 April15, 2018.

[12] J. W. Kim, M. Salauddin, H. Cho, M. S. Rasel, ,J. Y. Park, "Electromagnetic energy harvester based on a finger trigger rotational gear module and an array of disc Halbach magnets," Applied Energy, vol. 250, no. May, pp. 776–785, 2019.

[13] Kroener, Michael, Ravindran, Shankar, Woias, Peter, "Variable reluctance harvester for applications in railroad monitoring," Journal of Physics Conference Series, vol. 476, pp: 2091, 2013.

[14] Y. Zhang, J. Cao, H. Zhu, Y. Lei, "Design, modeling and experimental verification of circular Halbach electromagnetic energy harvesting from bearing motion," Energy Conversion and Management, vol. 180, pp. 811–821, 2019, 2018.

[15] G. Akoun, J. Yonnet, "3D analytical calculation of the forces exerted between two cuboidal magnets," in IEEE Transactions on Magnetics, vol. 20, no. 5, pp. 1962-1964, September 1984.

[16] F. Bancel. "Magnetic nodes," Journal of Physics D: Applied Physics, vol. 32, no. 17, pp: 32-2155, 1999.

Bimodal Anxiety State Assessment Based on Electromyography and Electroencephalogram

Diancong Zhang
College of Engineering
Qufu Normal University
Rizhao,China
1732136565@qq.com

Dianguo Cao*
College of Engineering
Qufu Normal University
Rizhao, China
caodg0318@163.com

Jingyu Zhang
College of Engineering
Qufu Normal University
Rizhao, China
1322733491@qq.com

Jiashuai Wang
College of Engineering
Qufu Normal University
Rizhao,China

Jingxin Cao
College of Engineering
Qufu Normal University
Rizhao,China

Xi Chen
College of Engineering
Qufu Normal University
Rizhao,China

Abstract—**Anxiety is a common emotional state of patients in rehabilitation training, which will affect rehabilitation training. In the current research, there are problems such as single use of electromyography signal to quantify anxiety and low accuracy. A method of bimodal feature fusion of electromyography (EMG) and electroencephalography (EEG) is proposed to realize anxiety assessment. First, the anxiety induced by 16 healthy subjects was recorded by EMG and EEG, comprehensive features were extracted from the signals. Then the pattern recognition model fused the features of EMG and EEG, based on theories of genetic algorithm, particle swarm optimization and support vector machine, which was designed to anxiety assessment by adjusting the feature fusion coefficient. According to the anxiety grading standard, three anxiety states were selected as severe, moderate and mild. Using only EEG data and EMG data as the sample sets, the corresponding recognition accuracy rates are 71.54% and 67.98%. However, using EEG and EMG feature fusion data as the sample sets, the average recognition accuracy rate was 81.49%, which increased by 9.95%, 14.01% compare to using EEG or EMG data individually. In short, the method proposed in this study improves the accuracy of anxiety state recognition and helps to better study patient anxiety.**

Keywords—EMG, EEG, anxiety state recognition, rehabilitation training platform

I. Introduction

In recent years, domestic and foreign research institutions have conducted research on robot-assisted rehabilitation training technology. Most of the robot rehabilitation training technologies focused on the patient's physical state monitoring and ignored the patient's psychological and emotional state monitoring, which may easily cause the training intensity to be inconsistent with the patient's mentality and lead to the training effect deterioration [1]. Anxiety is a common emotion in rehabilitation training. Although anxiety is an inner mental activity, it can still be expressed through facial expressions and limbs, such as frowning and unconscious body shaking.

However, it's not convincing that experimenter's feelings are not obvious when quantitatively assessing different degrees of anxiety. In response to this problem, many hospitals and scientific research institutions have proposed to allow experimenters to fill in the Hamilton Anxiety Scale and Likert Scale [2]. The grading can reflect the psychology of the patient to a certain extent, but there is a problem that the experimenter fills it out incorrectly. With the development of physiological signal technology, researchers use skin electrical, electrocardiogram, respiration, myoelectric and other physiological signals to study patients' anxiety recognition. Adheena M.A.[3] used electrocardiography to quantify individual anxiety responses. K. Mistry [4] proposed micro genetic (GA) embedded particle swarm optimization (PSO) optimization features to realize seven facial expression recognition. Asma Baghdadi[5] applied the classifier to classify anger and anxiety states with a rate of 67.72%. Mary Juliana[6] focused on the exhaustive analysis of EEG signal concerning anxiety quantification by k-nearest neighbors, linear discriminant analysis and quadratic discriminant analysis classifiers. Zheng[7] used EEG to analyze anxiety in cycling, and used principal component analysis and K nearest neighbors to classify the three-level anxiety recognition rate to 62.5%. Tian Man[8] calculated sample entropy and used Extreme Learning Machine to realize emotion recognition and classification.

In view of the above-mentioned problems of only recognizing anxiety without quantifying the anxiety level, too small difference in the adjacency of anxiety level classification, and low recognition rate of single use of EEG to assess anxiety, this article proposes a dual EMG and EEG-based Evaluation method of modal anxiety state. The design of experiment and data processing is elaborated in Section II. The classifier model is established in Section III. Through comparative analysis and simulation in Section IV, the benefit of bimodal anxiety state assessment method based on EMG and EEG is proved. Finally, the conclusion are provided in Section V.

II. Data

A. Participants

A total of 16 university students (24 ± 2, male: 10, female: 6) participated as volunteers. All participants had normal vision and hearing. They had no obvious emotional problems within 24 hours before the experimental test. In addition, experimenters need to understand the experiment process in advance and sign an informed consent form.

B. Experimental Method

The EMG collection equipment is the Delsys wireless surface EMG test system, and the collection locations(as shown in Fig. 1) are facial corrugator muscle (Ch1) and zygomatic major muscle (Ch2). The EEG acquisition equipment is 64-channel BP wireless portable EEG analyzer. The brain area related to emotions is the frontal lobe area. Only 8 channels such as FP1, FP2, FC5, FC6, F7, F8, FT9 and FT10 are used. Sampling frequency is 1000 Hz. The rehabilitation training platform is the Fourier M2 upper limb

* Dianguo Cao is corresponding author(e-mail: caodg0318@163.com).

rehabilitation robot platform, which mainly has passive, active, and impedance working modes. Fill in the Likert scale before the experiment and measure the signal. During the test, train in impedance, active, and passive modes in turn. Each mode was trained for 10 minutes, during which time rested for 2 minutes and filled out the anxiety scale. Each experimenter trains 25 groups, the training interval is more than one hour. The specific flow chart of the collection experiment is shown in Fig. 2.
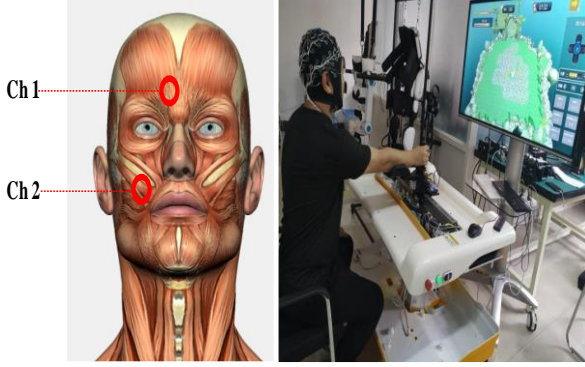
Fig.1. Experimental equipment and collection location.Ch1: Corrugator Muscle;Ch2: Zygomatic Major Muscle.
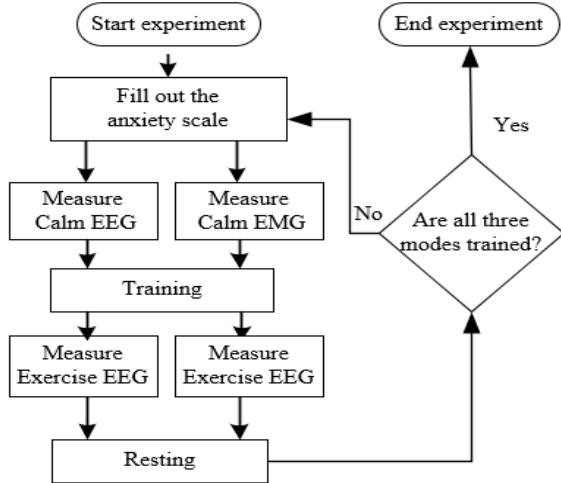
Fig.2. Experimental acquisition flowchart

C. Data Preprocessing

The EMG signal is a non-stationary signal with a weak amplitude and is susceptible to interference from external noise, so it is necessary to preprocess the collected EMG(as shown in Fig. 3). Original EMG signals were first filter out high-frequency noise interference through a 0-400Hz low-pass filter. Power frequency interference is removed through 50Hz adaptive notch filter. Finally, the ensemble empirical mode decomposition (EEMD) method is used to remove the baseline drift.
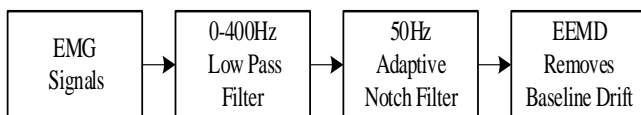
Fig.3. EMG signal preprocessing flowchart

The EEG signal amplitude is about 10-100 microvolts, and it is susceptible to interference from ocular electricity, myoelectricity and instruments during the acquisition process. EEG signal preprocessing mainly uses EEGLAB software to

perform band-pass filtering of 0.1-45 Hz, while eliminating power frequency interference. Independent component analysis is a useful statistical method for signal separation. Therefore, it was employed for manually separating artifacts from EEG signals.

D. Feature Extraction

Previous research results show that the frequency domain features can effectively analyze the transient performance of the signal. The frequency domain features selected in this paper are the average power frequency (MPF) and the median frequency (MF) [8].,which were calculated as follows:

$$EMG_{MPF} = \int_{f_{mf}}^{+\infty} fP(f)df \Big/ \int_0^{+\infty} P(f)df \qquad (1)$$

$$EMG_{MF} = \int_{f_{mf}}^{+\infty} P(f)df = \frac{1}{2}\int_0^{+\infty} P(f)df \qquad (2)$$

where $P(f)$ is the power spectral density, f_{mf} is the median frequency, and f is the frequency.

Previous research results show that entropy features can effectively distinguish different anxiety states [8][9][10]. The main entropy features selected in this paper are sample entropy, approximate entropy and fuzzy entropy.

Sample entropy (SampEn) is an improved time series complexity test method based on approximate entropy, and the higher the degree of brain confusion, the smaller the sample entropy [8]. SampEn is calculated as

$$SampEn(m,r,N) = -\ln\left[\frac{B^{m+1}(r)}{B^m(r)}\right] \qquad (3)$$

where m is dimension of the EMG signal, r is the similar tolerance, N is the length of the muscle signal. In this study, set $m=2$ and $r=0.25*\sigma$, where σ is the standard deviation of the EMG signal.

Fuzzy entropy (FuzzyEn) is less dependent on the event sequence, the transition is smoother when the parameters change, and the robustness to noisy signals is better [9]. FuzzyEn is calculated as

$$FuzzyEn(m,r,N) = \ln\varphi^m(r) - \ln\varphi^{m+1}(r) \qquad (4)$$

Approximate entropy (ApEn) is used to quantify the regularity and unpredictability of time series fluctuations, and represents the complexity of time series. Approximate entropy can be used to reflect the complexity of EEG signals [10].ApEn is calculated as

$$ApEn(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r) \qquad (5)$$

III. CLASSIFICATION MODEL

Considering that there is a certain correlation between EEG and EMG, this paper uses feature fusion to connect the feature information of EMG and EEG signals. The working principle is as follows: First, using the weighted average method to define the fusion coefficient to represent the trust degree and weight coefficient of the EEG and EMG feature vectors [12], then form the brain EMG feature fusion vector. Using SVM to learn and train the constructed feature fusion

vector, combining genetic algorithm and particle swarm algorithm to continuously optimize the feature fusion coefficient, find the optimal fusion coefficient, then import the test set into the classifier to check the classification effect.

A. Construct Fusion Feature Vector

We define $\lambda_i = \left[\lambda_{i1}, \cdots, \lambda_{ip} \right]$, $\eta_i = \left[\eta_{i1}, \cdots, \eta_{iq} \right]$ ($i = 1, \cdots, n$) as the feature vector EMG and EEG samples respectively, where n is number of samples, p, q are the characteristic dimensions of EMG and EEG. Suppose feature fusion coefficient matrix is $T = \left[t_1 \lambda_1, \cdots, t_n \lambda_n \right]$, then define the feature fusion vector of EEG and EMG as

$$x_i = \left[t_1 \lambda_{i1}, \cdots, t_p \lambda_{ip}, t_{p+1} \eta_1, \cdots, t_{p+q} \eta_{iq} \right] \quad (6)$$

B. Classifier

One-versus-rest (OVR) and support vector machine methods are combined to construct three classifiers. Suppose training set $\mathbf{X}_T$ has d samples, $\mathbf{V}_T$ is the training set label, then,

$$\begin{cases} \mathbf{X}_T = \left[x_1, x_2, \cdots, x_d \right]^T \\ \mathbf{V}_T = \left[v_1, v_2, \cdots, v_d \right]^T, v_i \in \{1, 2, 3\} \end{cases} \quad (7)$$

where $v_i = 1$ is mild anxiety, $v_i = 2$ is moderate anxiety, $v_i = 3$ is severe anxiety.

Since the radial basis kernel function has strong anti-interference ability against noise and there is only one kernel parameter, the radial basis kernel function is selected as the kernel function, then the corresponding classification decision function is:

$$f(x) = sign \left[\sum_{i=1}^{d} \alpha_i^* v_i \exp(-\frac{\|h_i\text{-}h_j\|^2}{2\sigma^2}) + b^* \right] \quad (8)$$

The advantages of particle swarm algorithm are simple and easy to implement, with storage and evolution capabilities and few adjustment parameters. However, there are defects that the discrete optimization problem is not handled well and it is easy to fall into the local optimization problem. The genetic algorithm has the advantages of good convergence, the global optimization has a good effect, high robustness and easy to combine with other algorithms. The disadvantage is that the programming is complicated and the parameter design is more difficult. Therefore, the GA-PSO hybrid algorithm [13] is formed by combining the advantages of particle swarm optimization and genetic algorithm in global and local optimization to optimize parameters.

The obtained optimal fusion coefficient matrix is combined with the sample feature vector to form a feature fusion matrix, which is input into the SVM classifier, and combined with the Likert scale to achieve anxiety state classification. The flowchart based on GA-PSO-SVM feature fusion and anxiety state classification is shown in Fig. 4.
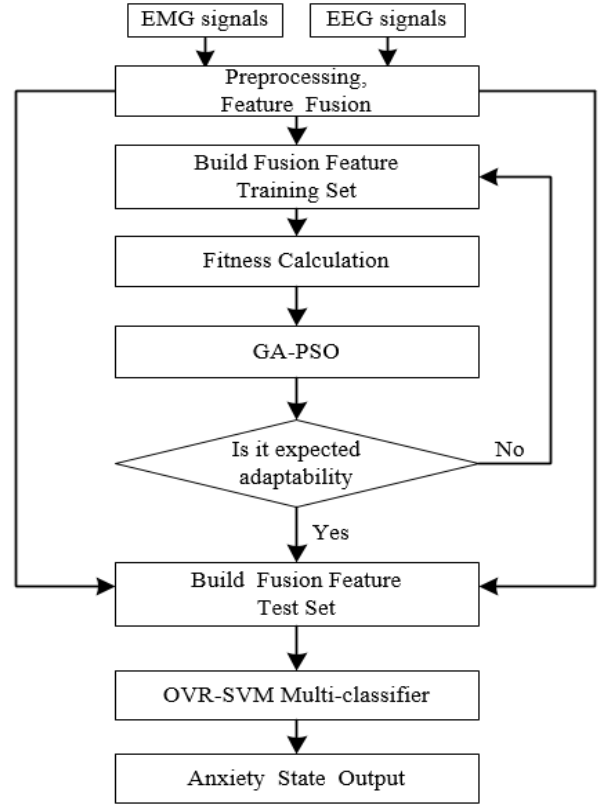
Fig.4. Anxiety State Classification Model Based on GA-PSO-SVM Feature Fusion

C. Classification Evaluation Index

Commonly used classification evaluation indicators include accuracy, precision, recall, f1_score, confusion matrix, ROC curve, etc. The confusion matrix, also known as the error matrix [14], is the most basic, most intuitive, and easiest method to calculate the accuracy of the classification model. It mainly calculates the precision (P), recall (R), and accuracy (ACC). Precision definition is defined as follows:

$$ACC_i = \frac{2 * P_i * R_i}{P_i + R_i}, i = 1, 2, 3 \quad (9)$$

where P_i is the precision of the i-th sample, and R_i is the recall of the i-th sample.

IV. RESULT ANALYSIS

A. Data Preprocessing

The spectrograms of the EMG signals denoising process are shown in Fig. 5. Fig. 5a shows the frequency spectrum of the original signal, there can be seen a lot of noise. Fig. 5b shows the frequency spectrum after filtering out high-frequency noise with a low-pass filter. The frequency spectrum after the 50Hz power frequency interference is removed by the adaptive notch filter is shown in the Fig. 5c is the frequency spectrum for removing power frequency interference. The spectrogram after removing the baseline drift by the EEMD method is shown in the Fig. 5d. Fig. 6 shows the EEG signal preprocessed by EEGLAB software. It can be seen from the figure that the noise such as eye electricity is basically removed, and the waveform is relatively smooth.
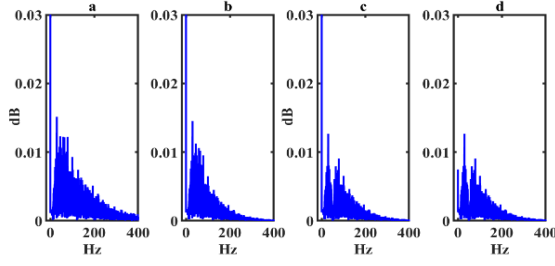
Fig.5. Spectrogram of EMG signal denoising process
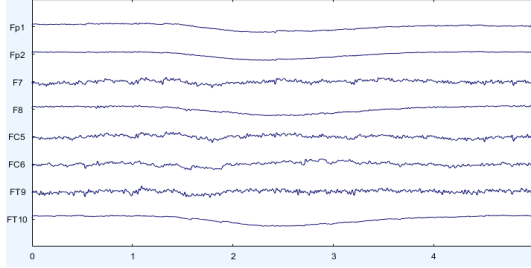
Fig.6. EEG filtering results

B. Analysis of Characteristics of Brain and Myoelectric Physiological Signals in Different Anxiety States

Fig. 7 shows the change trend of EMG frequency domain characteristics of an experimenter under different anxiety states. Fig. 7a and Fig. 7b selected features as MF and MPF respectively. As can be seen from the figure, the frequency domain characteristics of the overall three states have a good separation. However, there is overlap in adjacent areas of anxiety.
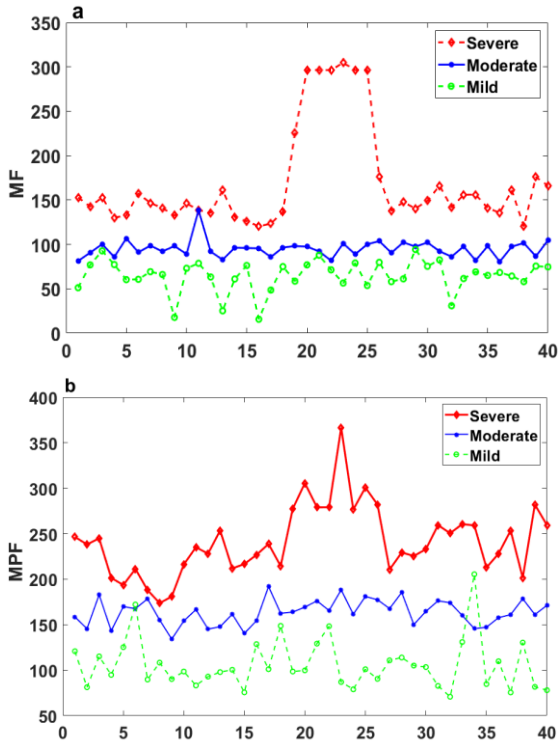
Fig.7. The experimenter's EMG frequency domain characteristic curve diagram under different anxiety states.

Fig. 8 is a comparison of the different entropies of the collected EEG signals in different training modes. The EEG signal of the first channel (FP1) after preprocessing is selected. From the figure, it can be seen that the approximate entropy and sample entropy are in different training modes The change

trend of fuzzy entropy is consistent and relatively obvious. Although the change of fuzzy entropy is not very obvious, the change trend is consistent, indicating that three kinds of entropy can be used as the characteristics of EEG signals.
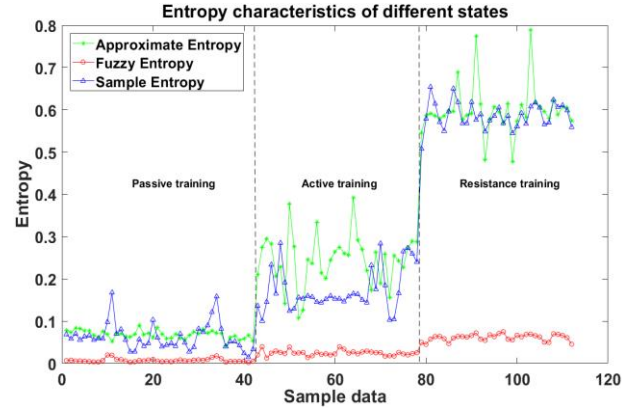
Fig.8. Comparison of different entropy features of EEG signals collected in different training modes.

C. Fusion Coefficient Optimization Process

The fusion coefficient is the key to construct the feature fusion vector of EMG and EEG signals, and determines the performance of the classifier. To prevent particle swarm from falling into local optimum, genetic algorithm and particle swarm algorithm are combined. Set the initial population size of PSO to 20, learning factor, inertia weight to 0.9, and mutation probability to 0.5. From the experimental data, 300 groups are selected for each state, a total of 900 groups. Selecting 600 sets of data as the training set to achieve iterative training through the support vector machine classifier. Using the other 300 groups as the test set to verify the performance of the classifier until it reaches the expected fitness of 95%. Fig. 9 is a diagram of the fitness convergence process. It can be seen that the number of iterations converges around 100 times.
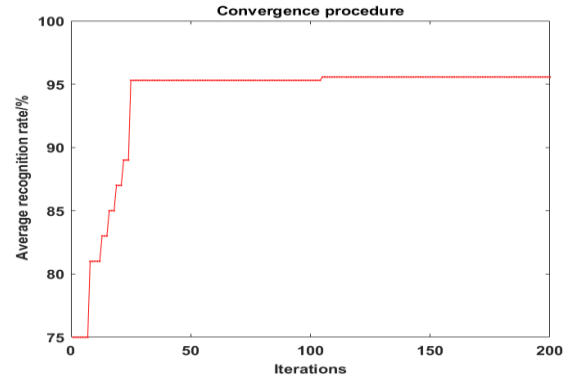
Fig.9. Fitness convergence process diagram

D. Different Methods to Identify Anxiety States

Table I shows the classification results of brain muscle signal fusion features by the improved GA-PSO-SVM classifier in this paper. The samples of each experimenter are used as sample sets for training and learning, then other samples are randomly selected for identification and classification. Table I is a sample of 8 people randomly selected from the experimenters. Each person randomly selects 600 sets of sample data as the training set and 100 sets of data as the test set. The classifier designed by this method has 84.80%, 76.06%, 83.61% recognition accuracy for severe, moderate, mild, and three anxiety states, and the average recognition accuracy rate is 81.49%.

Anxiety State		Experimenter ID								ACC
		1	2	4	7	9	10	13	16	
A1	P	81	76	83	77	81	77	86	76	84.80
	R	93	85	94	81	93	88	93	86	
A2	P	75	77	86	73	77	71	78	73	76.06
	R	75	76	82	77	76	71	76	74	
A3	P	93	85	90	80	90	88	93	86	83.61
	R	84	76	80	75	83	76	83	80	

In order to verify that the brain and EMG feature fusion has a better recognition effect than solely relying on EEG or EMG feature fusion, the EEG samples and EMG samples are used as training sets and the improved GA-PSO-SVM classifier is used for retraining.

TABLE II. EEG Features Recognition Accuracy Result (Unit: %)

Anxiety State		Experimenter ID								ACC
		1	2	4	7	9	10	13	16	
A1	P	68	70	75	71	67	70	74	68	72.25
	R	74	78	79	73	73	76	74	76	
A2	P	68	66	71	69	69	66	70	70	68.56
	R	70	65	72	68	66	66	70	71	
A3	P	73	77	78	73	76	76	79	79	73.80
	R	72	71	73	71	71	71	69	73	

TABLE III. EMG Features Recognition Accuracy Result (Unit: %)

Anxiety State		Experimenter ID								ACC
		1	2	4	7	9	10	13	16	
A1	P	71	64	70	61	60	75	64	65	69.57
	R	70	74	80	68	71	75	68	71	
A2	P	65	63	66	59	61	62	65	60	63.31
	R	62	63	68	65	61	65	68	60	
A3	P	73	75	80	72	73	70	70	73	71.06
	R	68	71	72	68	70	66	70	67	

Test and classification results are shown in Table II and Table III. Compared with Table I, it can be seen that the recognition results of EEG or EMG alone are 9.95% and 14.01% different from the results of brain and EMG fusion, indicating that the method proposed in this paper has a great effect on improving the recognition rate of anxiety states.

V. Conclusion

Aiming at excessive physical state detection and ignorance of anxiety and other mental state problems in the process of robot-assisted rehabilitation, a bimodal anxiety state assessment method based on EEG and EMG is proposed. Extracting the time and frequency characteristics of the EMG signal and the entropy characteristics of the EEG signal during the training of healthy experimenters. In order to find the optimal integration factor, using genetic algorithm and particle swarm optimization convergence coefficient. Using OVR and support vector machine to form a multi-classifier to realize anxiety state recognition. According to the anxiety grading standard, this paper selects three anxiety states as severe, moderate and mild. Only EEG data and EMG data are used as sample sets, the corresponding recognition accuracy rates are 71.54% and 67.98%. The corresponding recognition accuracy rates of EEG and EMG feature fusion classification were 84.80%, 76.06%, 83.61%, the average recognition accuracy rate was 81.49%. The results show that this study effectively improves the accuracy of anxiety assessment. This research is helpful to study the problem of interaction between patients and rehabilitation robot training, adjust training tasks according to the actual situation of patients and improve the efficiency of rehabilitation.

References

[1] G. Z. Xu, A. G. Song, X. Gao, S. Chen and B. G. Xu, "Robot-assisted clinical rehabilitation experiment based on anxiety emotion and hybrid control," Chinese Journal of Scientific Instrument,2017,pp. 2364-2372.

[2] X. M. Tang and X. G. Zang, "The feasibility of pupil determination to evaluate anxiety in patient with arrhythmia," China Modern Medicine, 2014, pp. 4-8.

[3] A. M. Adheena, N. Sindhu and S. Jerritta, "Physiological etection of Anxiety," 2018 International Conference on Circuits and Systems in Digital Enterprise Technology, Kottayam, India, 2018, pp. 1-5.

[4] K. Mistry, L. Zhang, S. C. Neoh, C. P. Lim and B. Fielding, "A Micro-GA Embedded PSO Feature Selection Approach to Intelligent Facial Emotion Recognition," IEEE Transactions on Cybernetics, vol. 47, no. 6, pp. 1496-1509, June 2017.

[5] A. Baghdadi and Y. Aribi, "Effectiveness of dominance for Anxiety Vs Anger detection," 2019 Fifth International Conference on Advances in Biomedical Engineering, Tripoli, Lebanon, 2019, pp. 1-4.

[6] M. J. M. Juliana and N. Rangaswamy, "A Neoteric Approach Towards Anxiety Quantification by EEG Signal Analysis," International Conference on Power, Energy, Control and Transmission Systems, Chennai, 2018, pp. 51-59.

[7] Y. Zheng, T. C. H. Wong, B. H. K. Leung and C. C. Y. Poon, "Unobtrusive and Multimodal Wearable Sensing to Quantify Anxiety," IEEE Sensors Journal, vol. 16, no. 10, pp. 3689-3696, May15, 2016.

[8] M. Tian, F. L. Yang and Y. Zhang, "Research on emotion recognition with EEG signal feature selection based on entropy," Electronic Measurement Technology, 2018, pp. 106-111.

[9] Z. Z. Luo, Z. H. Yan and W. D. Fu, "Electroencephalogram Artifact Filtering Method of Single Channel EEG Based on CEEMDAN-ICA," Chinese Journal of Sensors and Actuators,2018,pp. 75-80.

[10] D. L. Cai,Q. H. Zhang and Y. S. Zhu, "EEG emotion recognition using convolutional neural network with 3D input," Computer Engineering and Applications,in press.

[11] Z. Cao and C. Lin, "Inherent Fuzzy Entropy for the Improvement of EEG Complexity Evaluation,"in IEEE Transactions on Fuzzy Systems, vol. 26, no. 2, pp. 1032-1035, April 2018.

[12] P. Xie, Y. Y. Chen and Y. B. Hao, "Multimodal Fusion of EEG and EMG Signals for a Hybrid BCI. Research on hybrid brain-computer interface based on eeg fusion," in Chinese Journal of Biomedical Engineering, 2016, pp.20-30.

[13] Q. Wang and H. Y. Zhang, "Pedestrian Detection Based on GA-PSO Optimized Support Vector Machine," in Measurement & Control Technology,2019,pp.51-55.

[14] A. Schwarz, J. Pereira, R. Kobler and G. R. Müller-Putz, "Unimanual and Bimanual Reach-and-Grasp Actions Can Be Decoded From Human EEG," in IEEE Transactions on Biomedical Engineering, vol. 67, no. 6, pp. 1684-1695, June 2020.

Performance Degradation Prediction of Rolling Bearing based on KJADE and Holt–Winters

Ran Wei
College of Electrical Engineering and Automation, Anhui University
Hefei, China
wei_ran2019@163.com

Yawei Hu
College of Electrical Engineering and Automation, Anhui University
Hefei, China
yhu@ahu.edu.cn

Changbo He
College of Electrical Engineering and Automation, Anhui University
Hefei, China
changbh@ahu.edu.cn

Zheng Cao
College of Electrical Engineering and Automation, Anhui University
Hefei, China
caozheng@ahu.edu.cn

Siliang Lu
College of Electrical Engineering and Automation, Anhui University
National Engineering Laboratory of Energy-Saving Motor and Control Technology, Anhui University
Hefei, China
lusliang@mail.ustc.edu.cn

Fang Liu
College of Electrical Engineering and Automation, Anhui University
National Engineering Laboratory of Energy-Saving Motor and Control Technology, Anhui University
Hefei, China
ufun@ahu.edu.cn

Yongbin Liu*
College of Electrical Engineering and Automation, Anhui University.
National Engineering Laboratory of Energy-Saving Motor and Control Technology, Anhui University
Hefei, China
lyb@ustc.edu.cn

Abstract—**A performance degradation prediction method is proposed in this paper for condition monitoring and bearings performance degradation prediction. This method is the combination of kernel joint approximate diagonalization of eigen-matrices (KJADE) and Holt–Winters. First, the vibration signals acquired from running bearing are processed through multi-domain features extraction. An optimal feature set was obtained from the multi-domain features through dimensionality reduction and feature fusion using the KJADE algorithm. Then, the between- and within-class scatters were calculated to acquire the performance degradation indicators. Finally, the performance degradation prediction model based on Holt–Winters was established to predict the bearing performance degradation. Results show that bearing degradation trend can be effectively identified by the proposed method. Moreover, the prediction accuracy of this method is higher than that of extreme learning machine (ELM).**

Keywords—rolling bearing, performance degradation prediction, KJADE, Holt–Winters

I. INTRODUCTION

Prognostics Health Management (PHM) can realize the prediction of equipment failure before its occurrence and combine various available resource information to provide a series of maintenance guarantee measures, including diagnosis, prevention of equipment failures, reliability, and remaining useful life prediction of equipment's key components [1]. Industrial economic losses and human casualties can be avoided by accurately evaluating the bearing deterioration status of machine and finding bearing failure timely [2]. In the bearings performance degradation evaluation, the critical problems that must be solved include the performance degradation characteristics based on real-time signals and accurate prediction of the degradation [3].

In current studies, the single domain features are used as indicators of bearing degradation, such as root mean square (RMS) [4] or kurtosis [5]. Relative RMS, which was unaffected by individual bearing differences, was used to evaluate the regularity of bearing performance degradation [7]. The assessment model was established with combination of lifting wavelet packet decomposition and fuzzy c-means [6]. The RMS and kurtosis value were used as degradation indicators [8]. The time-domain statistical features of vibration signals, which were calculated by multi-sensor information fusion, were used to form fusion vectors for classification [9]. Multiple time-domain statistical characteristics were extracted to obtain the original feature and inputted into the stacked autoencoder network to obtain bottleneck feature for the bearing degradation prediction through the long short-term memory network [10]. Although the above research can obtain bearing degradation feature, comprehensively assessing the operation state of bearing using the characteristics from the separative domain is difficult. Therefore, the time and frequency domains feature must be fused. Such fusion aims to extract sensitive features and reduce redundancy of original data.

A method is proposed for bearing degradation prediction based on the combination of KJADE [11] and Holt–Winters. First, vibration signals acquired during the bearing operation are preprocessed, mixed-domain features set are extracted from the preprocessed signals. Next, feature dimension reduction is performed following the KJADE method to obtain the optimal feature parameter set. The between- and within-class scatters (SS) is also calculated from the optimal parameters to acquire the performance degradation indicator. Finally, the established perfor-

---

*Yongbin Liu is the corresponding author. (e-mail: lyb@ustc.edu.cn).

mance degradation prediction model is combined with the Holt–Winters to predict the bearing degradation.

II. FRAMEWORK OF BEARING PERFORMANCE DEGRADATION PREDICTION

The main idea of the bearing performance degradation prediction is to obtain the effective features that can characterize the bearing state through the signal needs to be preprocessed and feature extraction. And a certain method is used to assess bearing degradation status. This method can provide a criteria for predictive maintenance and prevent the occurrence of unexpected accidents.

The prediction process of bearing performance degradation in this paper is shown in Fig. 1. First, vibration signals acquired during the bearing operation are collected by the sensor, and mixed-domain features are extracted from the pre-processed data. The feature parameters set are normalized considering the different dimensions of the feature parameters. Then, the KJADE algorithm based on kernel function is used for feature fusion to obtain the optimal feature parameter set. Finally, the SS is calculated from the feature parameter set to acquire the degradation indicator, and Holt–Winters model is established to predict the bearing performance degradation process.
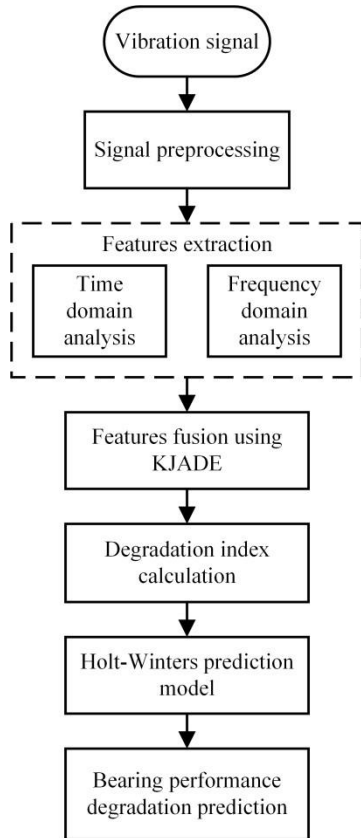
Fig. 1. Flow chart of bearing performance degradation prediction

III. OPTIMAL FEATURE EXTRACTION BASED ON KJADE

KJADE is a new feature fusion method, which combine kernel function with the JADE algorithm. The potential feature vector can be extracted through the kernel function in the high-dimensional feature space. And the JADE algorithm can extract feature which could reflect the bearing health status. KJADE has better nonlinear processing capabilities for bearing vibration signals.

KJADE is to map the observation data $X = \{x_1, x_2, ..., x_m\}$ to a high-dimensional feature space

F through a nonlinear function Φ, and the mapped feature space is $F = \{\Phi(x_1), \Phi(x_2), ..., \Phi(x_M)\}$. The JADE can then be used in this feature space to change the nonlinear separable problem into the linear one in F. The kernel function is introduced to obtain a kernel matrix K as follows:

$$K_{ij} = k(x_i, x_j) = \langle \Phi(x_i) \cdot \Phi(x_j) \rangle, \tag{1}$$

where x_i and x_j are the input vectors of sample space.

is required to meet the Mercer condition. The kernel functions include the following:

$$k(x_i, x_j) = (\alpha x_i^T x_j + c)^d , \tag{2}$$

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \tag{3}$$

$$k(x_i, x_j) = \tanh(\alpha x_i^T x_j + c) , \tag{4}$$

where the aforementioned kernel functions are respectively polynomial, Gaussian, and sigmoid kernel functions. Therefore, $k(x_i, x_j)$ is the inner product of two vectors in feature space F. In this paper, because Gaussian kernel function has high effectiveness in solving practical problems, so it is used. The matrix K is defined as Eq. (5).

$$K_{ij} = K_{ij} - \frac{1}{M}\sum_{r=1}^{M} K_{ir} - \frac{1}{M}\sum_{r=1}^{M} K_{rj} + \frac{1}{M^2}\sum_{r,s=1}^{M} K_{rs} . \tag{5}$$

In addition, the JADE is used to detect the hidden nonlinear relations in the observation signal. .

IV. PREDICTION MODEL BASED ON HOLT–WINTERS Holt–

Winters model is a practical approach to predict in some fields and the prediction model is built from an exponentially weighted average of past observed data [12]. The main idea of Holt–Winters is to combine with exponential smoothing method to evaluate the long-term trend, periodic change and trend increment. Then, a forecast model is established to extrapolate predictive value.

A time series $\{y_t\} = \{Y_1, Y_2, ..., Y_t\}$, $t = 1, 2, ..., m$ with trends and periodic effects is available. In the additive Holt–Winters model, three components of the time series are horizontal, trend, and periodic term, and their update formulas are respectively presented as follows:

$$L_t = \alpha(Y_t - S_{t-p}) + (1-\alpha)(L_{t-1} + T_{t-1}), \tag{6}$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} , \tag{7}$$

$$S_t = \gamma(T_t - L_t) + (1-\gamma)S_{t-p} , \tag{8}$$

where the parameters α, β, γ are horizontal, trend, and periodic smoothing coefficients, respectively, and their value range is $0 \le \alpha \le 1$ $0 \le \beta \le 1$, $0 \le \lambda \le 1$. p is the period length. Y_t, L_t, T_t, S_t are respectively the observed value, the stable component, the predominant trend component, and the periodic component of the time series at time t.

The model prediction formula is as follows:

$$F_{t+k} = L_t + kT_t + S_{t+k-p} , \qquad (9)$$

where   is the predicted points.

Holt–Winters method has a good effect on processing time series containing trends and periodicities for the bearing degradation features. So, the performance degradation indicator is calculated from the optimal feature set in Part II, and a prediction model is established in combination with Holt–Winters to predict the bearing degradation process.

V.   CASE ANALYSIS

The experiment uses four HRB6305 bearings. They were fixed on the same shaft and connected with the motor. A radial load of 750 kg was applied to all bearings to accelerate the bearing damage process, and the bearing speed was 3000 rpm. Full-life vibration signals were obtained by the NI PXI acquisition system. The vibration signals acquisition frequency was 20 kHz, the data were collected every 5 min. The experimental platform   is shown in Fig. 2.
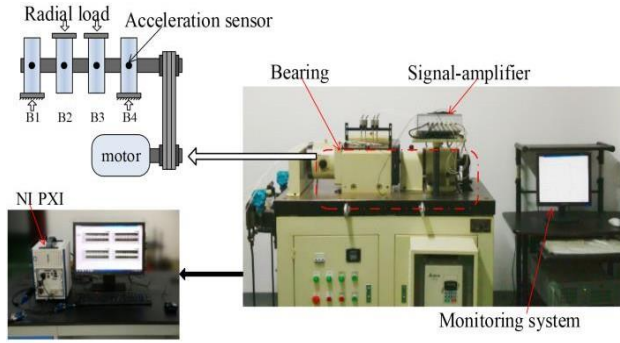
Fig. 2. Experimental setup

The fault of the rolling element is selected as the experimental object. And the full-life vibration signal data is shown in Fig. 3. First, the mixed-domain features are extracted from the preprocessed data. Then, KJADE is used for feature fusion to obtain optimal feature parameter set, and the SS is calculated from fusion features to acquire the degradation indicator. Finally, the established performance degradation prediction model is combined with the Holt–Winters to predict the bearing degradation. The first 10 groups of degradation indicators are used as input to obtain the initial value of the model, and the single- and multi-step predictions are performed. The error is measured by the RMSE. The predict results are shown in Fig. 4, and TABLE I. presents RMSE values.
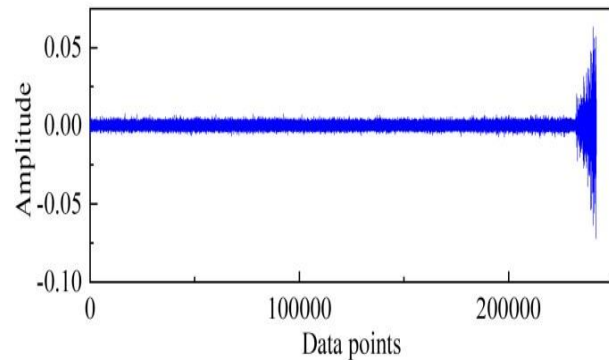
Fig. 3. Full-life vibration signal

application on the evaluation of bearing degradation status, and the accuracy is relatively high.
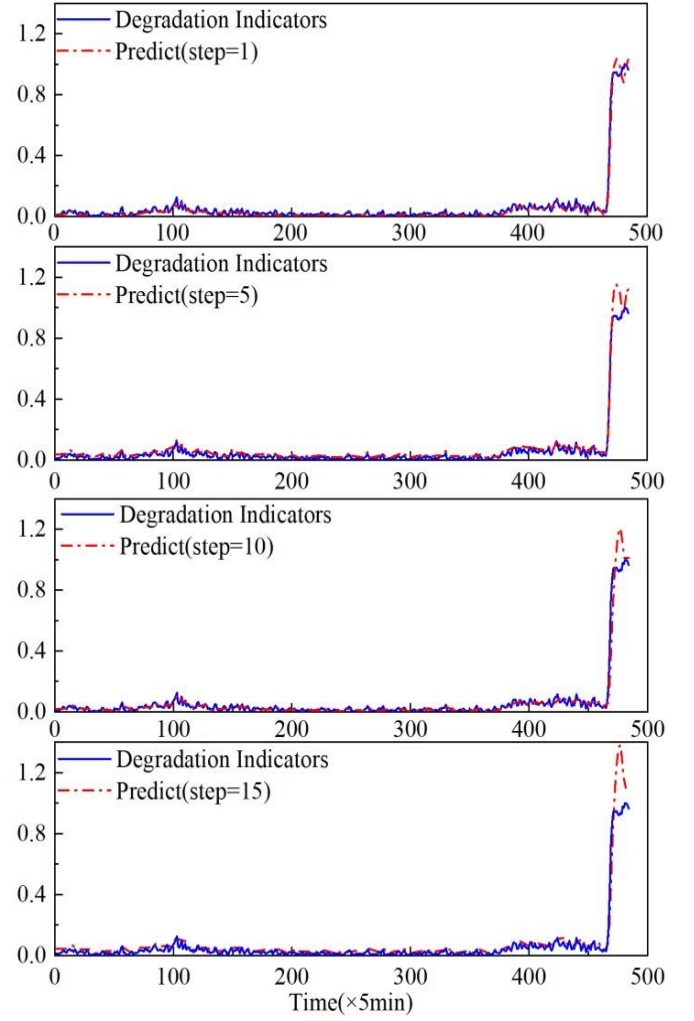
Fig. 4. Performance degradation prediction of the rolling bearings

The analysis of Fig. 4 and TABLE I. show that the prediction step is larger, the error is higher. The results show that the method can bring less error and higher pre- dicted precision. When the step is small, it has a good

TABLE I.   RMSE OF DIFFERENT PREDICTION STEP

Fault type	1	5	10	15
Rolling fault	0.0197	0.0294	0.0380	0.0537

Meanwhile, the ELM method is used to compare with the proposed method. Fig. 5 shows that when step = 10, the proposed method has smaller fluctuations in the prediction curve compared with the ELM method. These values are close to the true value, with small RMSE values as shown in TABLE II.

Fig. 5. Prediction result of the rolling bearings.

TABLE II. RMSE OF DIFFERENT PREDICTION STEP（STEP=10）

Method	The proposed method	ELM
RMSE	0.0400	0.0708

VI. CONCLUSION

In this paper, based on the combination of KJADE and Holt–Winters，a method is proposed for condition monitoring and degradation prediction of bearings. First, the vibration signal collected by the sensor is preprocessed. Multi-domain statistical feature parameters are also extracted to avoid the difficulty of comprehensively evaluating the bearing running state by single domain features. Second, the KJADE algorithm based on kernel function is used for feature fusion to obtain the optimal feature parameter set. The SS is calculated to acquire the performance degradation indicator, and the prediction model combined with Holt–Winters is established to predict the degradation process. This paper verifies that the proposed method can accurately identify degradation status, and have a good application on bearing degradation trend prediction through the analysis of bearing full life experiment. The prediction error of this method is also lower than that of the ELM.

REFERENCES

[1]. Lee, G. Y., et al. (2018). "Machine health management in smart factory: A review." Journal of Mechanical Science & Technology 32(3): 987-1009.

[2]. Wang, D., Zhao, Y., Yi, C., Tsui, K., Lin, J. (2018). "Sparsity guided empirical wavelet transform for fault diagnosis of rolling element bearings." Mechanical Systems and Signal Processing 101: 292-308.

[3]. Dong, W., Tsui, K.L., Qiang, M. (2018). "Prognostics and HealthManagement: A Review of Vibration based Bearing and Gear Health Indicators." IEEE Access, 6, 665–676.

[4]. Lei, Y., Li, N., Lin, J. (2016). "A New Method Based on StochasticProcess Models for Machine Remaining Useful Life Prediction." IEEE Transactions on Instrumentation & Measurement 65(12): 2671-2684.

[5]. Jing，X., Shen, C., Shi, J., Zhu, Z. (2018). Initial center frequency-guided VMD for fault diagnosis of rotating machines." Journal of Sound and Vibration 435: 36-55.

[6]. Pan, Y., Chen, J., Li, X. (2010). "Bearing performance degradationassessment based on lifting wavelet packet decomposition and fuzzy c-means." Mechanical Systems & Signal Processing 24(2): 559-566.

[7]. Chen, X., Shen, Z., He, Z., Sun, C., Liu, Z. (2013). "Remaining lifeprognostics of rolling bearing based on relative features and multi- variable support vector machine." Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 227(12): 2849-2860.

[8]. Elforjani, M. and S. Shanbr (2018). "Prognosis of Bearing AcousticEmission Signals Using Supervised Machine Learning." IEEE Transactions on Industrial Electronics 65(7): 5864-5871.

[9]. Jiang, L.L., Yin, H.K., Li, X.J., Tang, S.W. "Fault Diagnosis of Rotating Machinery Based on Multisensor Information Fusion Using SVM and Time-Domain Features." Shock and Vibration 2014(1) (2014) 153–154.

[10]. Tang, G., Zhou, Y., Wang, H., Li, G. (2018). "Prediction of bearing performance degradation with bottleneck feature based on LSTM network. " IEEE Instrumentation and Measurement Society.

[11]. Liu, Y., He, B., Liu, F., Lu, S., Zhao, Y. (2016). "Feature fusion using kernel joint approximate diagonalization of eigen-matrices for rolling bearing fault identification." Journal of Sound & Vibration 385: 389-401.

[12]. Sudheer, G. and A. Suseelatha (2015). "Short term load forecasting using wavelet transform combined with Holt–Winters and weighted nearest neighbor models." International Journal of Electrical Power & Energy Systems 64: 340-346.

Fault Diagnosis of Wind Turbine Bearing Using Variational Nonlinear Chirp Mode Decomposition and Order Analysis

Zhichao Wu
College of Electri. Eng. Automa.
Anhui University
Hefei, China
496968987@qq.com

Jiahao Niu
College of Electri. Eng. Automa.
Anhui University
Hefei, China
niu.jiahao@qq.com

Siliang Lu*
Traction Power State Key Laboratory
Southwest Jiaotong University
Chengdu, China
lusliang@mail.ustc.edu.cn

Yongbin Liu
College of Electrical Engineering
and Automation
Anhui University
Hefei, China
lyb@ustc.edu.cn

Fang Liu
College of Electrical Engineering
and Automation
Anhui University
Hefei, China
ufun@foxmail.com

Abstract—In this paper, a method for bearing fault diagnosis without using a tachometer under variable speed conditions by combining the variational nonlinear chirp mode decomposition (VNCMD) with angular resampling technique is proposed. First of all, the original bearing vibration signal can be decomposed and the rotation mode can be reconstructed by using the VNCMD. Then the rotation phase is calculated from the reconstructed rotation mode. In the next step, the original bearing vibration signal is resampled on the basis of the extracted rotation angle information. Finally, the bearing fault indicator can be displayed in the envelope order spectrum and then the bearing fault can be diagnosed. This method addresses the problem that the frequency components are aliasing in the envelope spectrum under variable speed conditions. The reliability and effectiveness of the method proposed in this study are proved by the experiments.

Keywords—bearing fault diagnosis; VNCMD; angular resampling; order analysis

I. INTRODUCTION

Wind power generation is a resource with great potential and a new way to use clean energy efficiently. It has the advantages of low composition, clean and pollution-free, short construction period and small floor space. For coastal islands, grassland and pastoral areas, mountainous areas and plateau areas where water and fuel are scarce and transportation is inconvenient, wind power generation is very suitable for 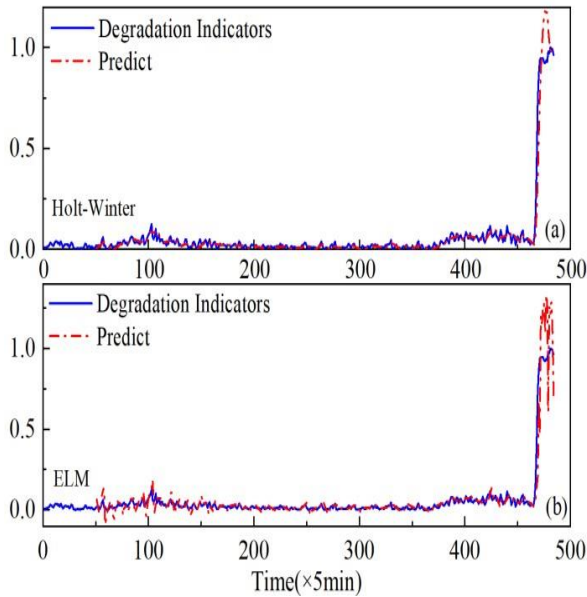local conditions. China is rich in wind energy resources, so it is of great significance to study wind power generation to save resources. The wind turbine blades are whirled by the wind, which converts the wind energy into mechanical energy and then into electricity. The wind is a natural energy source that cannot be controlled by human beings. The size and direction of the wind often change, which makes the speed unstable and the wind turbines are becoming complex, large and intelligent, so the problem of bearing fault often occurs in wind turbines. The severe bearing fault will lead to economic loss and even catastrophic. With the development of signal processing technology, people have higher and higher requirements for bearing fault diagnosis technology. Although the traditional state monitoring

method can realize effectively fault diagnosis, it has some limitations. Under this circumstance, condition monitoring and fault diagnosis of wind turbine bearings are significant and necessary in industry [1]. It is the main problem to select the proper fault diagnosis method of wind turbine.

The tested direct-drive wind turbine is a permanent magnet synchronous generator (PMSG) which is driven by a permanent magnet synchronous motor (PMSM) under variable speed. PSMG composed of stator and rotor has simple structure, small volume, small loss and high efficiency. The design of PMSM makes it more efficient at capturing wind energy. Therefore, it is widely used in aerospace, wind power generation and other fields. Permanent magnet is used to replace electric excitation in PMSG. The stator of PMSG is composed of the magnetizing stator core and the conducting stator winding as well as some components used for fixing the core and winding. There is a pair of magnetic poles with excitation windings placed on the rotor core. The excitation device is not needed and the whole structure is simplified. PMSG can be divided into tangential, radial, axial and hybrid structure of rotor magnetic circuit according to the relationship between the magnetization direction of permanent magnets and the rotation direction of rotor. When the excitation winding of PMSG is on direct current, the excitation magnetic field will be formed between the generator air gap, and then making the armature winding generate induction electromotive to drive the motor operation.

When the rotation speed of the bearing of wind turbines keep changing, the frequency of the vibration signal is changing at every moment, which leads to a phenomenon that the frequency components are smearing in the spectrum, and it is difficult to diagnosis the kind of bearing fault accurately. Order analysis (OA) method has been proposed to solve this problem. This technology has developed rapidly in recent years. OA method is the conversion of an equal time sampling interval signal into an equal angle sampling signal and then analyze the spectrum of the angle sampling signal. There are two types of OA methods, one is based on a tachometer and the other one can be realized without using a tachometer. The latter is called as tacholess OA (TOA) [2]. In practice, it is costly and complex to install the tachometers. Given this, TOA methods were proposed in order to figure out this problem by extracting the rotational speed information from the

---

original bearing vibration signal [3]. After extracting the rotational information, the original bearing vibration signal in time-domain can be resampled in angular-domain. The advantage of this method is that it requires no hardware.

TOA can be realized by using the time-frequency analysis (TFA) methods and the signal decomposition methods. TFA is now a major research hotspot in the field of signal analysis. In order to obtain the time-varying spectrum characteristics of signals, many scholars proposed various forms of time-frequency distribution functions. The celebrated TFA methods include wavelet transform [4], short-time Fourier transform (STFT) [5], envelope demodulation method and Hilbert-Huang transform (HHT) [6]. In fact, the STFT is calculated by dividing a longer time signal into shorter segments of the same length, and computing the Fourier transform on each of the shorter segments. The latter two methods are now widely used. The classical signal decomposition methods include empirical mode decomposition (EMD) [7], and variational mode decomposition (VMD). EMD can adaptively decompose a nonlinear frequency-modulation signal into the signal with different frequency characteristics in time domain [8],[9],[10]. VMD method was proposed to analyze the nonlinear chirp signals [11]. The original bearing vibration signal is decomposed into some intrinsic mode functions (IMFs) according to the frequency characteristics by using EMD or VMD method. Each IMF is a stationary component and contains the local characteristics of the original signal on different time scales. A multi-component signal can be decomposed into a fixed number of IMFs that basically fluctuate around their central frequencies by using the VMD method. But if the number of decomposed signals is too large, especially in the high frequencies, the signals will be discontinuous and not very regular. The VMD method has its limits that it can only works for narrow-band signals. Therefore, an analysis method suitable for wider range of signals is urgently needed.

In recent years, an adaptive signal analysis method called variational nonlinear chirp mode decomposition (VNCMD) was always proposed to analyze the wide-band nonlinear chirp signals [12] and was used in various industries. A wide-band signal converts into a narrow-band signal by using the demodulation technic of VNCMD method, and then the decomposition problem is converted to find an optimal solution of modulation. The original signal is decomposed into several nonlinear chirp modes (NCMs) according to the estimated instantaneous frequency (IF) curves. Then the IFs and decomposed modes are updated iteratively until the narrowest bandwidth signal is obtained.

Given this, the VNCMD method can be introduced to make fault diagnosis by analyzing the original bearing vibration signal. This method firstly extracts the rotation information, and then the bearing vibration signal is resampled by TOA. The original vibration signal is decomposed into several NCMs according to the higher frequency of energy after the time-frequency analysis, and each mode can represent a kind of feature of the original vibration signal, and then the rotation mode can be reconstructed. The rotation phase of the bearing shaft is computed according to the reconstructed rotation mode. Finally, the original vibration signal is resampled by using

resampling technique and the envelope order spectrum of the resampled signal is computed to make the diagnosis of the bearing fault type.

The rest of the paper is organized as follows. Sec. II introduces the fault bearing diagnosis method proposed based on VNCMD and TOA. Sec. III introduces the experimental setup of the proposed method. Sec. IV and Sec. V introduce the experimental results and conclusions, respectively.

II. METHODS

The proposed method can be considered to be a two-step process: 1) the original vibration signal is decomposed and the rotation mode is reconstructed by using the VNCMD, and then the rotational speed information is extracted; and 2) the original bearing vibration signal is resampled by resampling technique according to the rotation phase extracted from rotational speed information, and then the envelope order spectrum is calculated. It becomes easy to identify the kind of bearing fault.

A. Reconstructing Rotation Mode by VNCMD and Calculating Rotation Phase

The original bearing vibration signal is represented as $V(t)$ and it is a multicomponent signal that can be decomposed into k components as:

$$
\begin{aligned}
V(t) &= \sum_{i=1}^{k} A_i(t)\cos\left(2\pi\int_0^t \widetilde{f}_i(s)ds\right) + N(t) \\
&= \sum_{i=1}^{k} m_i(t)\cos\left(2\pi\int_0^t \widetilde{f}_i(s)ds\right) \\
&\quad + \sum_{i=1}^{k} n_i(t)\sin\left(2\pi\int_0^t \widetilde{f}_i(s)ds\right) + N(t),
\end{aligned}
\tag{1}
$$

in which $A_i(t)$ is the instantaneous amplitude (IA) of the ith signal component. $N(t)$ is a noise component generated during measurement. $f_i(t)$ is the IF of the ith signal component. There are two demodulated signals separately called $m_i(t)$ and $n_i(t)$ that is described as follows:

$$
m_i(t) = A_i(t)\cos\left(2\pi\int_0^t (f_i(s) - \widetilde{f}_i(s))ds\right), \tag{2}
$$

$$
n_i(t) = -A_i(t)\sin\left(2\pi\int_0^t (f_i(s) - \widetilde{f}_i(s))ds\right). \tag{3}
$$

First, the STFT of the original vibration signal is calculated as:

$$
STFT_V(t,f) = \int_{-\infty}^{+\infty} [V(u)\delta^*(u-t)]e^{-j2\pi fu}du. \tag{4}
$$

The frequency with the highest energy can be selected by STFT. A hypothesis can be made that the information of data is sampled at $t = t_0,...,t_{N-1}$, the problem is discretized as:

$$
\min_{\{m_i\},\{n_i\},\{f_i\}} \left\{ \sum_i (\|\Omega_{m_i}\|_2^2 + \|\Omega_{n_i}\|_2^2) \right\}
$$
$$
s.t. \quad \left\| V - \sum_i (A_i m_i + B_i n_i) \right\|_2 \le \varepsilon, \tag{5}
$$

in which Ω is a second-order difference operator as:

$$
\Omega = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -2 & 1 \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}. \tag{6}
$$

Next, two demodulated orthogonal signals m_i and n_i are calculated iteratively as:

$$m_i^{x+1} = \left(\frac{2}{\alpha} \Omega^T \Omega + A_i^T A_i \right)^{-1} A_i^T$$
$$\cdot \left(V - \sum_{m \neq i} X_i m_i - \sum_m Y_i n_i - \omega - \frac{1}{\alpha} \lambda \right), \qquad (7)$$

and

$$n_i^{x+1} = \left(\frac{2}{\alpha} \Omega^T \Omega + B_i^T B_i \right)^{-1} B_i^T \qquad (8)$$
$$\cdot \left(V - \sum_m X_i m_i - \sum_{m \neq i} Y_i n_i - \omega - \frac{1}{\alpha} \lambda \right),$$

in which ω is a noise variable. The reconstructed rotation mode is calculated by iterative algorithm as:

$$V_i^{x+1}(t) = X_i(t) m_i^{x+1}(t) + Y_i(t) n_i^{x+1}(t), \qquad (9)$$

in which

$$X_i = diag[\cos(\varphi_i(t_0))\ldots\cos(\varphi_i(t_{N-1}))], \qquad (10)$$

$$Y_i = diag[\sin(\varphi_i(t_0))\ldots\sin(\varphi_i(t_{N-1}))], \qquad (11)$$

in which $\varphi_i(t) = 2\pi \int_0^\pi f_i(s) ds$. x is the number of total iterations.

By using Hilbert transform (HT), the rotation phase is calculated according to the reconstructed rotation mode as follows:

$$\widehat{V}_i^{k+1}(t) = H[V_i^{x+1}(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{V_i^{x+1}(\tau)}{t - \tau} d\tau, \qquad (12)$$

in which $H(\cdot)$ represents HT, and the rotation phase is computed as follows:

$$\phi(t) = angle(\widehat{V}_i^{x+1}(t)), \qquad (13)$$

in which $angle(\cdot)$ is a function in order to extracting phase information from a multiple variable. The angle curve of the rotation shaft can be computed as:

$$d(t) = \frac{180°}{\pi} \times (unwrap(\varphi(t))) \cdot \qquad (14)$$

B. Bearing Fault Diagnosis

After the bearing rotation phase is extracted from the reconstructed rotation mode, the original bearing vibration signal can be resampled on the base of the rotation phase as follows:

$$V'(t) = resample(V(t), d(t)), \qquad (15)$$

in which $resample(\cdot)$ is a function that already exists in MATLAB. By using the function of MAYLAB, the envelope spectrum of $V'(t)$ can be computed after the signal is resampled. After that, the fault characteristic order (FCO) in the envelope order spectrum is identified to diagnose the bearing fault. The FCO compresses the first order of the axis rotation frequency in the traditional order spectrum into a small part of an FCO. The algorithm flow chart of the proposed signal analysis method is shown in Fig. 1.

III. EXPERIMENTAL SETUP

In order to verify the efficiency and the practicability of the method proposed in the paper, an experimental test rig is designed as shown in Fig. 2. The tested direct-drive wind turbine is a PMSG. The PMSG is driven by a PMSM under varying speed condition. A faulty bearing on the
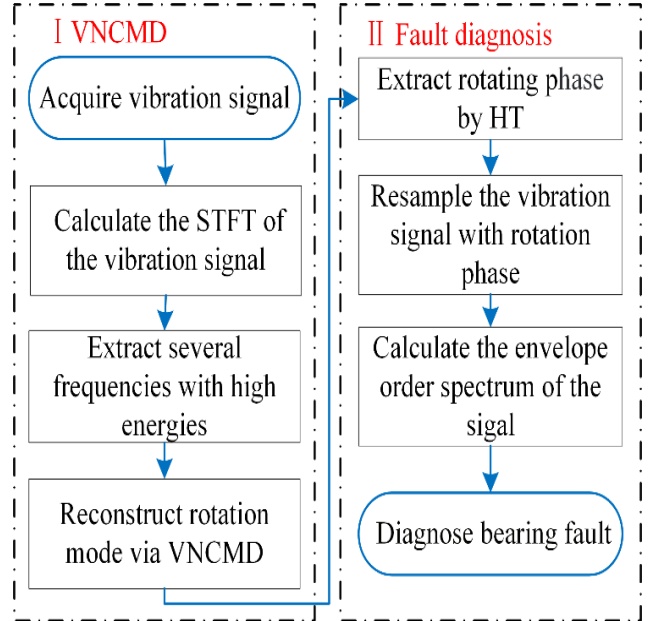
Fig. 1. Algorithm flow chart of the proposed signal analysis method.

outer raceway is installed on one-end of the PMSG, and the vibration signal is acquired by a piezoelectric acceleration sensor (CA-YD-1182, SINOCERA, Inc.). The signal is obtained by a data acquisition system (DAS, USB4432, NI, Inc.). The duration of the sampling process is 4s and sampling frequency of this experiment is set to 20 kHz, respectively. The FCO_o (denotes FCO at outer raceway) can be computed as:

$$FCOo = \frac{n}{2}\left(1 - \frac{d_1}{d_2} \cos \alpha \right), \qquad (16)$$

in which n is the quantity of rolling elements, d_1 is rolling element diameter, d_2 is pitch diameter, and α is contact angle. The parameters of the test PMSG and bearing are severally shown in Table I and Table II. The FCO_o is calculated as 3.055 by using (16).

TABLE I. PARAMETERS OF THE PMSG

Type	Output rated Voltage	Output rated power	No. of pole pairs	No. of phase
NE-100	24 VDC	100 W	6	3

TABLE II. PARAMETERS OF THE BEARING

Kind of bearing	External diameter	Internal diameter	No. of rollers	FCO_O
6203	40 mm	17 mm	8	3.055
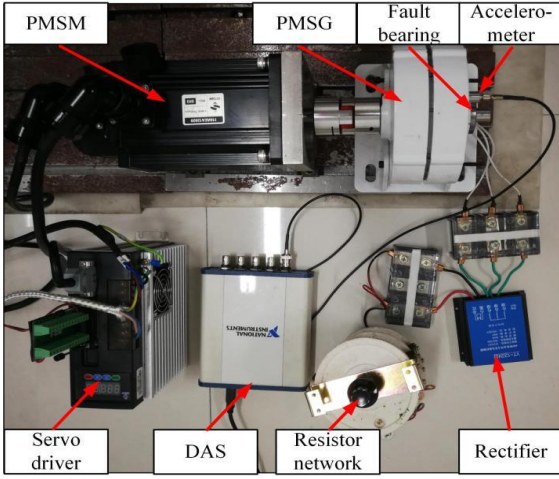
Fig. 2. Experimental setup.

IV. EXPERIMENAL RESULTS

The amplitude of the original bearing vibration signal gradually attenuates from 0 s to 4 s as is shown in Fig. 3(a). Some fault-induced pulses can be roughly observed in Fig. 3(a), but the intervals of the pulses are hardly identified. It is impossible to make fault diagnosis of the bearing due to other distractions. The spectrum of the vibration signal is shown in Fig. 3(b), where it can be discovered that the signal with higher energy locates at 0-4kHz, and the frequency components are overlapped with each other. The energy locates at higher than 4 kHz is almost zero. The vibration signal is demodulated using HT and the envelope spectrum is shown in Fig. 3(c). It can be found that the signal with higher energy locates at 0-50Hz. The energy locates at higher than 50Hz is almost zero and the frequency components are still overlapped with each other. It is difficult to make bearing fault diagnosis, because the envelope order spectrum doesn't have a fault indicator that related to the bearing fault type.

For the purpose of revealing the essential structure of the original vibration signal, the time-frequency representation of the original bearing vibration signal is computed by using STFT as shown in Fig. 4. The energy of the signal decays gradually from low frequencies to high frequencies. Multiple IF curves can be found in the time-frequency image, and the IFs fluctuate with the rotating speed. The higher the frequency, the greater the fluctuation and the smaller the energy. The energy of IFs concentrates at 0~150 Hz. Several IF curves can be clearly distinguished in Fig. 4, and the IF with the highest energy distributes at around 13~20 Hz as shown in Fig. 4. As a result, the frequency around 13~20 Hz is used to reconstruct the rotation mode.

After that, the original vibration signal is decomposed by using the VNCMD method and then the reconstructed modes of the vibration signal are extracted. The number of decomposed components is set as 2, 3, 4 and 5, separately. And the initial IF sequences are set as 13 Hz, 40 Hz, 60 Hz, 80 Hz, and 120Hz, severally. Through experimental verification, the performance is most significant when the decomposition level is set as 2. The initial IF sequences are set as 13 Hz and 40 Hz, severally. According to the experiment, the reconstructed mode 1 is associated with bearing rotation. When selecting other frequencies to reconstruct the rotational modes or other decomposed

quantities, FCO cannot be seen visually. Therefore, mode 1 is used as the final extracted rotation mode as shown in Fig. 5(a). The signal is clear and the noise interference can be neglected. It can be observed that the reconstructed rotation mode is similar to a sinusoidal signal. The rotation angle is calculated using HT and the result is shown in Fig. 5(b). It can be found that the angle curve increases with time with a nonlinear mode. It can be seen in Fig. 5(a) that the whole approximates a smooth ascending line.
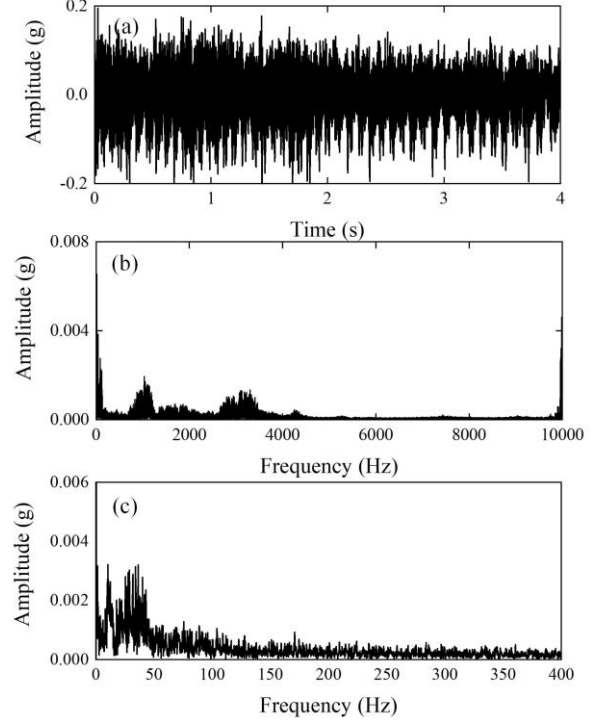
Fig. 3. The original bearing vibration signal: (a) signal waveform, (b) signal spectrum, and (c) signal envelope spectrum.
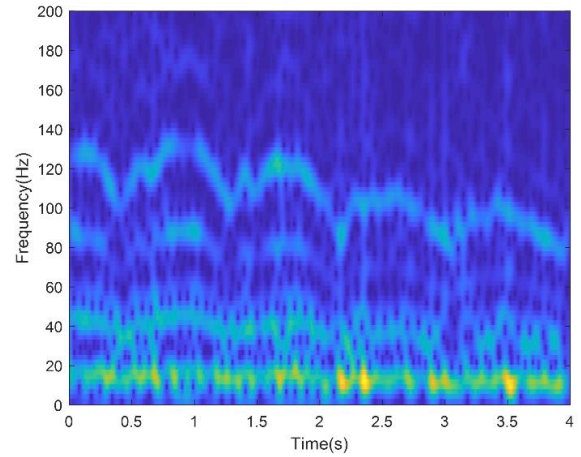
Fig. 4. STFT of the original bearing vibration signal.

Later, the original bearing vibration signal can be resampled on the basis of the calculated rotation angle, and the image of the resampled signal is shown in Fig. 6(a). Obviously, it is shown that the x-axis is changed from time to revolutions after signal resampling according to the calculated rotation angle. The amplitude of the signal gradually decreases with revolutions from 0 to 16. The resampling signal is more regular than the original equal-time increment sampled signal. Some of the interference

has been removed. But it is still unable to accurately diagnose the bearing fault. Then the spectrum is shown in Fig. 6(b). The higher amplitude is mainly concentrated in the lower orders. The envelope spectrum of the resampled signal is shown in Fig. 6(c).
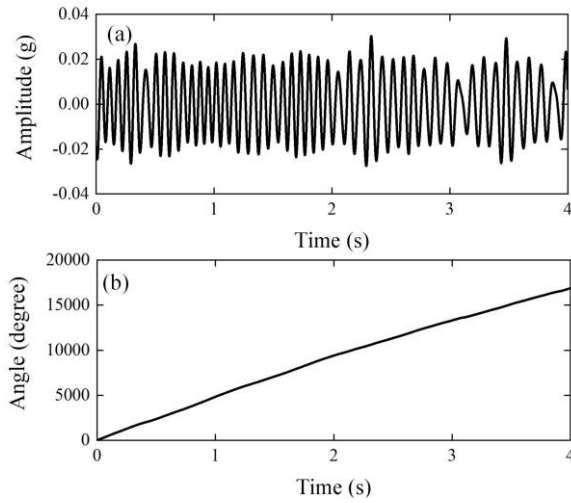
Fig. 5. (a) The reconstructed rotation mode, and (b) the calculated rotation angle curve.

It is visually to observe the fault indicators including FCO_O and its $2\times$ and $3\times$ harmonics, as indicated by the arrows in the envelope order spectrum without interference. It is clear that the FCO_o is about 3 in Fig. 6(c), and the second and third harmonic of FCO_o is about 6 and 9, respectively. These results approximate the value calculated earlier and demonstrate that there is an outer raceway fault with the bearing in the experiment. In the meanwhile, the pre-set bearing fault in the experimental setup has been confirmed. The rotation speed estimation and bearing fault diagnosis can be implemented by only processing the vibration signal in this study, and the proposed method can be used for varying-speed bearing fault diagnosis without using a tachometer. The accuracy and the convenience of this method is proved to be higher in the experiments.

V. CONCLUSIONS

In this study, a bearing fault diagnosis method of wind turbine called TOA is proposed. It can be used when the rotation speed of bearing changes with time. First, the original bearing vibration signal of which sampling frequency is fixed is obtained by a data acquisition system. After that, the time-frequency information of the time-varying bearing vibration signal is analyzed by using a time-frequency analysis method like STFT. And the time-frequency information calculated from the original bearing vibration signal is clearly displayed in the time-frequency image. According to the IF curve with high energy in the time-frequency image, the signal is decomposed and the rotation mode is reconstructed by using the VNCMD method. The curve of the rotation angle of the bearing is computed by using the HT and unwrapping operation. On the basis of the calculated rotation angle, the original bearing vibration signal is then resampled. Finally, the envelope order spectrum of the

resampled signal is calculated. After that, the bearing fault is diagnosed according to the fault indicator FCO and its harmonics. The method described in this study is used for bearing fault diagnosis without a tachometer, which provides considerable convenience. This method can be used when a fault of a wind turbine bearing happens without using a tachometer under variable speed conditions.
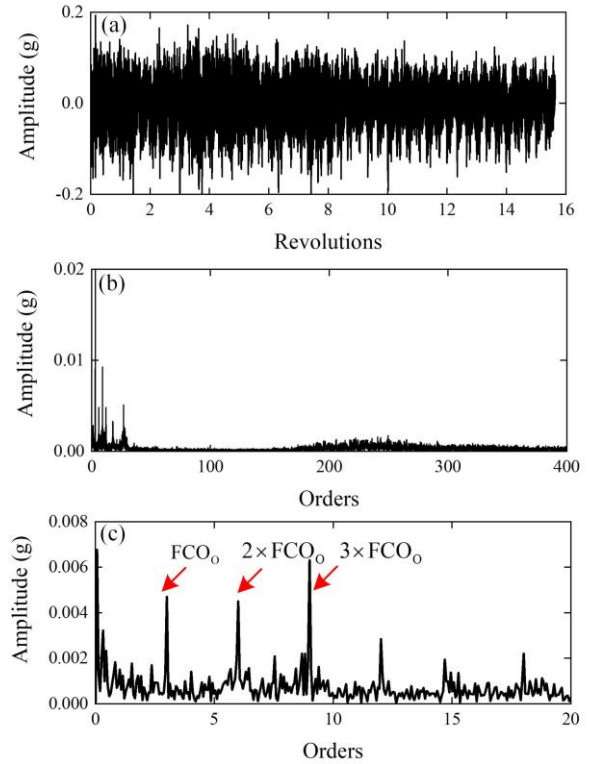
Fig. 6. The resampled signal: (a) signal waveform, (b) spectrum, and (c) envelope order spectrum.

REFERENCES

[1] Guo, J., Lu, S., Zhai, C., & He, Q. (2018). Automatic bearing fault diagnosis of permanent magnet synchronous generators in wind turbines subjected to noise interference. Measurement Science and Technology, 29(2), 025002.

[2] Wang, Y., Xu, G., Luo, A., Liang, L., & Jiang, K. (2016). An online tacholess order tracking technique based on generalized demodulation for rolling bearing fault detection. Journal of sound and vibration, 367, 233-249.

[3] Shi, J., Liang, M., Necsulescu, D. S., & Guan, Y. (2016). Generalized stepwise demodulation transform and synchrosqueezing for time–frequency analysis and bearing fault diagnosis. Journal of Sound and Vibration, 368, 202-222.

[4] Li, J., Wang, H., Wang, X., & Zhang, Y. (2020). Rolling bearing fault diagnosis based on improved adaptive parameterless empirical wavelet transform and sparse denoising. Measurement, 152, 107392.

[5] Gao, H., Liang, L., Chen, X., & Xu, G. (2015). Feature extraction and recognition for rolling element bearing fault utilizing short-

time Fourier transform and non-negative matrix factorization. Chinese Journal of Mechanical Engineering, 28(1), 96-105.

[6] Li, Q., & Wang, H. (2013). A research review of Hilbert-Huang Transform used for rolling bearing fault diagnosis. In Applied Mechanics and Materials (Vol. 397, pp. 2152-2155). Trans Tech Publications Ltd.

[7] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 454(1971), 903-995.

[8] Zhou, F., Yang, L., Zhou, H., & Yang, L. (2016). Optimal averages for nonlinear signal decompositions—Another alternative for empirical mode decomposition. Signal Processing, 121, 17-29.

[9] Pustelnik, N., Borgnat, P., & Flandrin, P. (2014). Empirical mode decomposition revisited by multicomponent non-smooth convex optimization. Signal Processing, 102, 313-331.

[10] Hou, T. Y., & Shi, Z. (2011). Adaptive data analysis via sparse time-frequency representation. Advances in Adaptive Data Analysis, 3(01n02), 1-28.

[11] Dragomiretskiy, K., & Zosso, D. (2013). Variational mode decomposition. IEEE transactions on signal processing, 62(3), 531-544.

[12] Chen, S., Dong, X., Peng, Z., Zhang, W., & Meng, G. (2017). Nonlinear chirp mode decomposition: A variational method. IEEE Transactions on Signal Processing, 65(22), 6024-6037.

Compressed Sensing-based Blade Tip-timing Vibration Reconstruction under Variable Speeds

Hao Sheng
College of Electrical and Information Engineering
Hunan University of Technology
Zhuzhou, China
18856895159@163.com

Zhongsheng Chen*
College of Electrical and Information Engineering
Hunan University of Technology
Zhuzhou, China
chenzs_hut@sina.com

Yemei Xia
College of Electrical and Information Engineering
Hunan University of Technology
Zhuzhou, China
1162853220@qq.com

Weimin Wang
College of Mechanical and Electrical Engineering
Beijing University of Chemical Technology
Beijing, China
wwm@mail.buct.edu.cn

Abstract—**Blade tip-timing (BTT) has been regarded as a promising solution of on-line blade vibration monitoring. The rotating speed is often considered to be constant in traditional BTT methods. In practice, this assumption is hardly satisfied, so that BTT vibration monitoring under variable speeds faces is a big problem to be solved. Moreover, BTT vibration signals are always under-sampled due to the limited number of BTT probes and multi-band with less prior knowledge due to system's nonlinearity and complicated aerodynamic excitations. Thus blind multi- band vibration reconstruction under variable speeds is a key challenge by using under-sampled BTT signals. To deal with it, a novel compressed sensing (CS) method in angular domain is proposed to overcome the challenge in this paper. First, angular-domain sampling model of BTT signals is built and its multi-coset sampling scheme is first presented. Then the CS model of BTT signals is derived in order domain. Two metrics of the support reconstruction ratio and the relative root mean square are defined to characterize the reconstruction performance in order and angular domains, respectively. In next simulations, the performances of four reconstruction algorithms are compared, i.e., Orthogonal Matching Pursuit, Multiple Signal Classification, Modified Focal Under- determined System Solver and Basis Pursuit Denoising algorithms. Influences of different algorithms and measurement noises on the reconstruction performance are simulated.**

Keywords—Blade tip-timing, Multi-band vibration reconstruction, Compressed sensing, Sub-sampled

I. INTRODUCTION

High-speed rotating blades are very important components in modern turbo-machinery, such as compressor blades, turbine blades, fan blades, and so on. At the same time, rotating blades are often exposed to severe environments during operation, including strong vibrations, large centrifugal forces and thermal stresses. Thus, different kinds of faults or damages are often caused in rotating blades as a result of high or low cycle fatigue. More than 60% of the overall faults are due to vibrations according to statistic data. Furthermore, more than 70% of blade faults are induced by vibrations [1-3]. Thus, accurately on-line vibration monitoring of rotating blades is much required. As an on-line and non-intrusive way, blade tip-timing (BTT) has been proved to be an advanced method of blade vibration monitoring, which can overcome the shortcomings of strain gauges [4, 5]. Traditional BTT method has two intrinsic drawbacks. The first one is undersampling. BTT sampling frequency is equal to the product of the rotating frequency and the number of BTT sensors, which is often less than double mode frequencies. Thus the Nyquist sampling theorem does not hold, so the BTT signals are under-sampled. The second one is that rotating speeds should be constant, otherwise leading to inaccurate vibration measurements. While rotating speeds hardly keep constant during operations in reality due to aerodynamic load instability, variable conditions, and other dynamic factors.

As we all know that true blade vibration characteristics can hardly be recovered by using under-sampled BTT signals. Thus it is much necessary to study reconstruct algorithms. Salhi *et al.* presented a technique to reconstruct a continuous BTT signal [6]. Chen *et al.* proposed a Shannon theorem-based under-sampled blade vibration reconstruction method [7] and later wavelet packet transform was introduced to decrease reconstruction aliasing [8]. For these methods, however, the BTT signal is looked as a narrow-band signal and its central frequency should be known. Otherwise, they are not feasible again. In practice, however, blade vibrations tend to be multi- frequency or multi-band due to system's nonlinearity and complicated aerodynamic excitations. In particular, when incipient damages appear in rotating blades, super-/sub- harmonic responses or/and frequency modulations will happen [9, 10]. In this case, single-frequency analysis methods cannot be used for blind multi-band vibrations. In order to release these strict conditions, Lin *et al.* tried to use sparse representation for multi-mode blade vibration signal reconstruction without prior knowledge [11]. But this work was based on the assumption of constant speed.

For blade vibration signals, central positions and widths of multiple bands always change with rotating speeds and working conditions [9, 10]. Thus the band-pass sampling and reconstruction method can hardly be used for BTT signals under variable speeds [12]. In recent years, multi-coset sampling (MCS) has been proposed to decrease the sampling frequency of multi-band signals [13, 14], where multiple parallel channels with different time offsets are used to collect a multi-band signal using the same sampling frequency. Signal reconstruction of the MCS process is seen as a Multiple-Measurement Vectors (MMV) problem [15]. Furthermore, the MCS reconstruction process can be projected into a compressed sensing (CS) framework. The CS theorem was first proposed by Donoho *et al.* [16], which is a novel sampling pattern against the Nyquist sampling.

Up to now, CS has been widely used for cognitive radio [17], medical signal processing [18] and image processing [19], and so on. The CS theory takes fully advantages of sparsity and compressibility of signals, so that one can recover certain signals by using under-sampled signals, instead of Nyquist-sampled ones. To date, there are two main groups of CS reconstruction algorithms [20]: 1) greedy algorithms which find the dominant components of the solution; and 2) relaxation methods which solve convex and non-convex problems. In the last few years, CS began to attract much attention for BTT signals. Xu *et al.* first built a CS model of BTT measurements and reconstructed BTT signals by minimizing the ℓ_1 norm [21]. Pan *et al.* proposed a dictionary learning method for representing sparse BTT vibration signals and then reconstructed them by the basis pursuit algorithm [22]. Spada *et al.* discussed several application conditions for BTT data [23]. However, variable speeds have randomly been considered before.

In summary, blade vibrations under variable speeds can be looked as blind multi-band signals. Therefore, it is promising and innovative to use the MCS and CS theory to reconstruct blade vibrations from under-sampled BTT measurements. Under variable speeds, however, BTT sampling is not equally spaced in time domain, so that the MCS cannot be used directly. Fortunately, the BTT sampling is a natural angular sampling process, so that the problem due to variable speeds will be overcome in angular domain. Different from existing CS methods in time domain, the innovation of this paper is the first to use the MCS and CS theory to reconstruct blade vibrations under variable speeds by using BTT signals in angular domain.

II. Angular Sampling Model Of BTT Measurements

A. Basic Principles of the BTT Method

Basic principle of BTT method is schematically demonstrated in Fig.1. *I* BTT probes are mounted in a stationary casing around a bladed-disk including *K* blades. In addition, one once-per-revolution (OPR) sensor is placed in front of the shaft to generate timing reference signals. It is assumed that the OPR sensor is on the same radius as the first BTT probe and the OPR marker is on the same radius as the first blade. Then the angles of the $i^{th}(1 \le i \le I)$ BTT sensor and the $k^{th}(1 \le k \le K)$ blade are denoted as α_i and θ_k, respectively.

Measuring times of arrival (TOAs) of each blade passing each BTT probe is the basis of the BTT method. Without vibrations under normal conditions, theoretical TOAs of each blade are fixed as a function of its rotating speed, rotating radius and position angle. Thus TOAs can be calculated theoretically. While blade vibrations happen, the blade will pass each BTT probe either earlier or later than normal. In this case, actual TOAs will shift from theoretic values, leading to time differences, by which blade vibration displacements can be calculated.

The bladed-disk is considered to rotate clockwise and its rotating speed is constant. When no vibrations, the theoretic TOAs of the kth blade from the ith BTT probe is described as
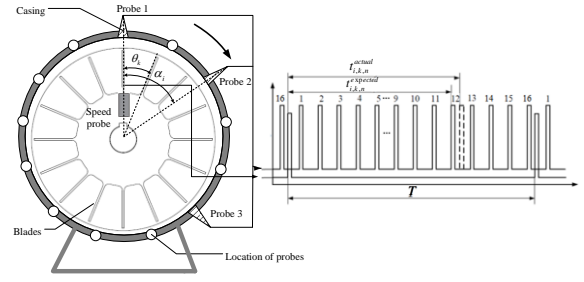
Fig. 1. Schematic diagram of the BTT method.

$$t_{i,k,n}^{expected} = \begin{cases} \dfrac{1}{2\pi f_r}\Big[2\pi(n-1)+\alpha_i-\theta_k\Big], \theta_k < \alpha_i \\ \dfrac{1}{2\pi f_r}\Big[2\pi n+\alpha_i-\theta_k\Big], \theta_k > \alpha_i \end{cases} \quad (1)$$

where f_r and n denote the rotating frequency and the nth revolution, respectively.

Then vibration displacements of the kth blade from the ith BTT probe can be calculated as,

$$d_{i,k}[n] = 2\pi f_r R(t_{i,k,n}^{expected} - t_{i,k,n}^{actual}) \quad (2)$$

where R denotes the rotating radius and $t_{i,k,n}^{actual}$ denotes the actual TOAs of the kth blade.

Based on Eq.(2), all-blade vibration displacements can be obtained simultaneously, which is also the biggest advantage of the BTT method compared with using strain gauges. However, once the rotating speed changes under actual operation conditions, two problems have to be faced. Firstly, actual TOAs ($t_{i,k,n}^{actual}$) become unknown and the rotating frequency f_r is not fixed, so that Eq.(2) cannot be used. Secondly, the sampling time interval is not a constant, leading to unequally-sampled vibration signals.

B. BTT Measurements in Angular Domain

As we all know that angular sampling has always been applied to deal with variable speeds [24]. Moreover, the angles of BTT probes in Fig.1 are fixed. Thus BTT sampling is a natural angular sampling processing, which is independent of rotating speed. Then we will derive blade vibration displacements in angular domain, instead of time domain.

Theoretical angles of the kth blade passing the ith BTT probe can be represented as,

$$\theta_{i,k,n}^{expected} = \begin{cases} 2\pi(n-1)+\alpha_i\text{-}\theta_k, (n=1,2,\cdots), \beta_k \le \alpha_i \\ 2\pi(n-1)+2\pi+\alpha_i\text{-}\theta_k, (n=1,2,\cdots), \beta_k > \alpha_i \end{cases} \quad (3)$$

Next, the rotating frequency of each revolution is assumed to a constant due to only using one OPR sensor. Then the average rotating frequency of each revolution is used for calculating actual angles of the kth blade from the ith BTT probe, which are shown as follows.

$$\theta_{i,k,n}^{actual} = 2\pi(n-1)+2\pi f_n \times \left(t_{i,k,n}^{actual}-\sum_{p=1}^{n-1}T_p\right) \quad (4)$$

where T_p, f_p are the average rotating period and frequency of the pth revolution, respectively.

Combining Eq.(3) and Eq.(4), angular-domain vibration displacements of the kth blade from the ith BTT probe can be calculated as,

$$y_i^k[n] = \left(\theta_{i,k,n}^{actual} - \theta_{i,k,n}^{\exp ected}\right) \times R \qquad (5)$$

C. MCS Scheme of BTT Vibrations in Angular Domain

True angular-domain vibration displacement of the kth blade is denoted as $y_k(\theta)$, which will be sampled for I times during each revolution according to Fig.1. Based on Eq.(3), the ideal impulsive sampling function can be represented as,

$$p_k^i(\theta) = \begin{cases} \sum_{n=-\infty}^{+\infty} \delta\left[\theta - (2\pi n + \alpha_i - \beta_k)\right], (n=0,1,2,\cdots), \beta_k \leq \alpha_i \\ \sum_{n=-\infty}^{+\infty} \delta\left[\theta - (2\pi n + 2\pi + \alpha_i - \beta_k)\right], (n=0,1,2,\cdots), \beta_k > \alpha_i \end{cases} \qquad (6)$$

where δ denotes the Dirac delta function. Then angular-domain vibration signals of the kth blade can be calculated as

$$\hat{y}_k(\theta) = y_k(\theta) \sum_{i=1}^{I} p_k^i(\theta) = \sum_{i=0}^{j-1} \sum_{n=-\infty}^{+\infty} y_k(\theta)\delta\left[\theta - (2\pi n + 2\pi + \alpha_i - \beta_k)\right]$$
$$+ \sum_{i=j}^{I-1} \sum_{n=-\infty}^{+\infty} y_k(\theta)\delta\left[\theta - (2\pi n + \alpha_i - \beta_k)\right] \qquad (7)$$

where $j = \left\{ j \mid \alpha_{j-1} \leq \beta_k \leq \alpha_j \right\}$.

Furthermore, We can see that such a sampling scheme can be looked as I-channel parallel sampling with fixed angular delay. Obviously, angular sampling scheme of BTT signals is equivalent to a MCS framework. Then the MCS theorem can be used here [13].

For any blade, the following steps are the same, so the subscript 'k' will be omitted for the sake of brevity. The maximum vibration Engine Order (EO) of the blade is assumed to be $E_o^{\max}$. L BTT probes are assumed to uniformly mount in the casing and the number of each probe is marked from 1 to L clockwise. Here L is an odd integer. According to the Nyquist sampling theorem, L should be selected to satisfy $L \geq 2E_o^{\max}$ in order to perfectly reconstruct vibration EOs of the blade. Then the Nyquist-sampled vibration signal in angular domain can be represented as,

$$y[n] = y(n\Delta\theta), \Delta\theta = 2\pi/L \qquad (8)$$

In practice, the I BTT probes can be considered to be chosen from the above L BTT probes, which is called as a MCS pattern (L, I, C), where the set $C = \{c_i : 1 \leq i \leq I\}$, c_i is the order number of the ith BTT probe and $1 \leq c_1 < c_2 \cdots < c_I \leq L$.

In this case, angular-domain sampling vibration signals from the ith BTT probe can be represented as,

$$\bar{y}_i[n] = y[nL + c_i], i = 1, \ldots, I \qquad (9)$$

where the sampling angle interval of each BTT probe is denoted as $\Delta\theta_I = L\Delta\theta$. Furthermore, the MCS scheme of BTT signals in angular domain can be plotted as Fig.2.
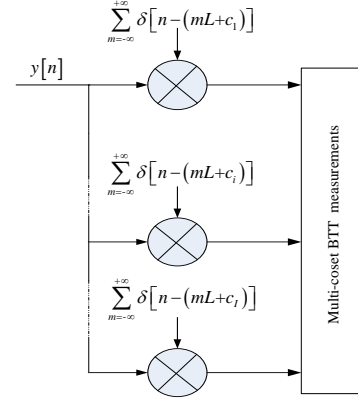
Fig. 2. Multi-coset sampling scheme of BTT signals in angular domain.

III. COMPRESSED SENSING MODEL OF BTT VIBRATION SIGNALS IN ORDER DOMAIN

In practice, I has to be selected so that $I \ll L$ due to the restrictions of space and cost. It is not difficult to understand that the BTT vibration measurements are also under-sampled in angular domain. That is to say, vibration EOs of the blade cannot be perfectly reconstructed by directly using $\bar{y}_i[n]$. However, we can seek another idea. If we can reconstruct Nyquist-sampled $y[n]$ by using $\bar{y}_i[n]$, true vibration features of the blade will also be obtained. In order to achieve it, we will build the relation between $y[n]$ and $\bar{y}_i[n]$ in order domain as follows.

Firstly, we will define an intermediate angular sampling signal as follows.

$$yy_i[n] = \begin{cases} y[n], n = mL + c_i, m \in Z \\ 0, \text{ otherwise} \end{cases} \qquad (10)$$

Discrete Fourier transform (DFT) of $yy[n]$ can be calculated as,

$$YY_i(O) = \sum_{n=-\infty}^{+\infty} yy_i[n]e^{-j2\pi On\Delta\theta} = \bar{Y}(O)e^{-j2\pi Oc_i\Delta\theta} \qquad (11)$$

where O denotes the order and $\bar{Y}(O)$ is the DFT of $\bar{y}[n]$.

Furthermore, the DFT of $yy_i[n]$ can also be calculated as,

$$YY_i(O) = \sum_{n=-\infty}^{+\infty} y[n] \frac{1}{L\Delta\theta} \sum_{k=0}^{L-1} e^{j\frac{2\pi}{L}k(n-c_i)} e^{-j2\pi On\Delta\theta}$$
$$= \sum_{k=0}^{L-1} e^{-j\frac{2\pi}{L}kc_i} Y\left(O - \frac{k}{2\pi}\right) \qquad (12)$$

where $O \in [-1/2L, 1/2L]$.

By combing Eq.(11) and Eq.(14), we will have

$$\bar{Y}_i(O) = e^{j2\pi Oc_i\Delta\theta} \sum_{k=0}^{L-1} e^{-j\frac{2\pi}{L}kc_i} Y\left(\left.O_0 - \left(k - \frac{L+1}{2}\right)\right/L\right) \qquad (13)$$

We can see from Eq. (13) that the entire order spectra of $Y(O)$ is equally divided into L narrow bands and the bandwidth is equal to $1/L\Delta\theta$, shown in Fig.3. Order spectra of the lth sub-band can be represented as follows.

$$Y_k(O) = Y\left(O_0 - \left(k - \frac{L+1}{2}\right)\middle/L\right), k = 1, 2, \cdots L \qquad (14)$$

where $O_0 = [-1/2L, 1/2L]$.
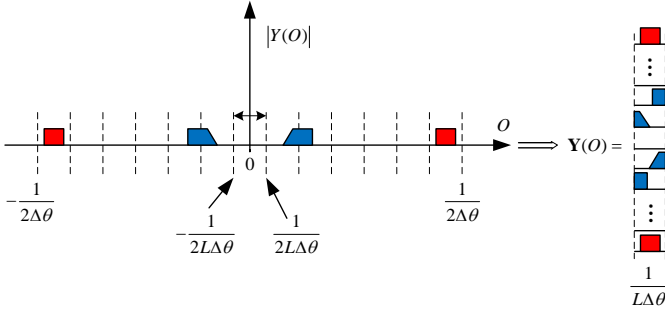
Fig. 3. The relation between $Y(O)$ and $Y_k(O)$.

For I BTT probes, Eq. (15) is further expressed in matrix form as follows.

$$\overline{\mathbf{Y}}(O) = \mathbf{\Phi}\mathbf{Y}(O) \qquad (15)$$

where,

$$\overline{\mathbf{Y}}(O) = [\overline{Y}_1(O)e^{-j2\pi O c_1\Delta\theta}, \cdots \overline{Y}_I(O)e^{-j2\pi O c_I\Delta\theta}]^T,$$
$$\mathbf{Y}(O) = [Y_1(O), \cdots Y_I(O)]^T,$$

$$\mathbf{\Phi} = \begin{bmatrix} e^{(-j2\pi c_1/L)} & \cdots & e^{(-j2\pi c_1 l/L)} & \cdots & e^{(-j2\pi c_1 L/L)} \\ \vdots & \vdots & \vdots & & \vdots \vdots \\ e^{(-j2\pi c_i/L)} & \cdots & e^{(-j2\pi c_i l/L)} & \cdots & e^{(-j2\pi c_i L/L)} \\ \vdots & \vdots & \vdots & & \vdots \vdots \\ e^{(-j2\pi c_I/L)} & \cdots & e^{(-j2\pi c_I l/L)} & \cdots & e^{(-j2\pi c_I L/L)} \end{bmatrix}.$$

In Eq.(17), $\overline{\mathbf{Y}}(O)$ can be obtained by I actual BTT measurements in angular domain. Thus the reconstruction solution is to calculate the unknown $\mathbf{Y}(O)$ using $\overline{\mathbf{Y}}(O)$ and $\mathbf{\Phi}$. Generally speaking, Eq.(17) is under-determined ($I \ll L$) so that $\mathbf{Y}(O)$ cannot be uniquely reconstructed. However, $\mathbf{Y}(O)$ is sparse and $Y_k(O)$ is its kth slice, so $\mathbf{Y}(O)$ only has few non-zero rows. In this case, the above blind multi-coset sampling of angular-domain BTT signals becomes the CS problem. Then the CS theory can provide a promising way to solve the reconstruction problem of unknown $\mathbf{Y}(O)$.

Now the CS model of angular-domain BTT vibration signals can be depicted as

$$\mathbf{P0}: \begin{array}{c} \min \|\mathbf{Y}(O)\|_0 \\ s.t. \ \overline{\mathbf{Y}}(O) = \mathbf{\Phi}\mathbf{Y}(O) \end{array} \qquad (16)$$

where $\|\cdot\|_0$ is the ℓ_0-norm.

As we all know that the solution of Eq.(16) is a NP-hard problem, so we have to find a near-optimal solution close to its exact solution. The common way is to replace the P0 problem by the P1 problem as follows.

$$\begin{array}{c} \min \|\mathbf{Y}(O)\|_1 \\ s.t. \ \overline{\mathbf{Y}}(O) = \mathbf{\Phi}\mathbf{Y}(O) \end{array} \qquad (17)$$

where $\|\cdot\|_1$ is the ℓ_1-norm.

Generally speaking, The P1 problem can be easily solved by the optimization theory. By now, many existing methods have been proposed to solve the P1 problem. In this paper, we consider the following four algorithms, including Orthogonal Matching Pursuit (OMP), Multiple Signal Classification (MUSIC), Modified Focal Underdetermined System Solver (MFOCUSS), and Basis Pursuit Denoising (BPDN).

IV. NUMERICAL SIMULATIONS

A. Simulated Multi-band Vibration Signals

In order to generate multi-band blade vibrations under variable speeds, the following multi-band signal $x(t)$ is simulated as

$$\begin{aligned} x(t) &= A_1 \sin(2\pi f_1 t + \varphi_1) + A_2 \sin(2\pi f_2 t + \varphi_2) \\ &+ A_3 \sin(2\pi f_3 t + \varphi_3) \end{aligned} \qquad (18)$$

where, f_1 and f_2 are used to simulate synchronous blade vibrations, f_3 denotes the frequency of one asynchronous vibration.

The rotating speed of each revolution is assumed to proportionally increase from r_L to r_H and f_1, f_2 are defined as follows.

$$f_i = EO_i\left(r_L + \frac{(n-1)K}{2}\right)/60 \qquad (19)$$

where, $K = (r_H - r_L)/N$, EO_i denotes the ith-order EO ($i = 1, 2$) and $n = 1, \ldots, N$.

In this paper, simulation parameters are set as: $r_L = 1000$ RPM, $r_H = 1250$ RPM, $EO_1 = 2$, $EO_2 = 3$, $A_1 = 1$ mm, $A_2 = 0.5$ mm, $A_3 = 1$ mm, $f_3 = 50$ Hz and $\varphi_1, \varphi_2, \varphi_3 = 0$.

Furthermore, two metrics are introduced to evaluate the performance of different reconstruction algorithms [25]. The first one is the support reconstruction ratio defined as follows.

$$R = \frac{Number\ of\ correct\ reconstructions}{Number\ of\ sample\ sets} \qquad (20)$$

The second one is the relative root mean square ($RRMS$) defined as follows.

$$RRMS = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}[i] - y[i])^2}{\sum_{i=1}^{n} y^2[i]}} \qquad (21)$$

where $y[i]$ and $\hat{y}[i]$ are the original and reconstructed multi-band signals, respectively. $RRMS$ can be considered as the reconstruction error in angular domain. Obviously, the signal tends to be perfectly reconstructed when $RRMS$ is close to zero or R is close to one.

In order to simulate the MCS process, the Nyquist sampling rate is chosen as $L = 13$. Then for any given sampling pattern (L, I, C), the corresponding simulated MCS samples can be generated based on Eq. (9). Here an example is given as follows, where the sampling pattern is (13, 4, [1, 2, 4, 10]). The signal-to-noise ratio is equal to 1dB and then the sampled signal in angular domain is shown in

Fig. 4.
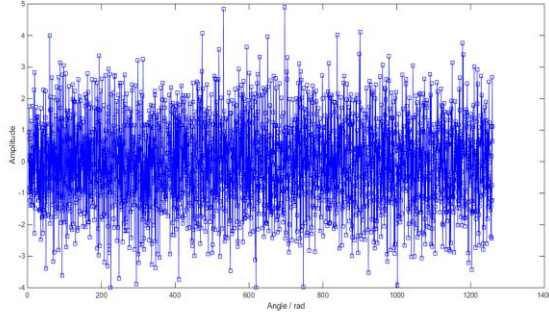
Fig. 4. Nyquist-sampled signals in angular domain.
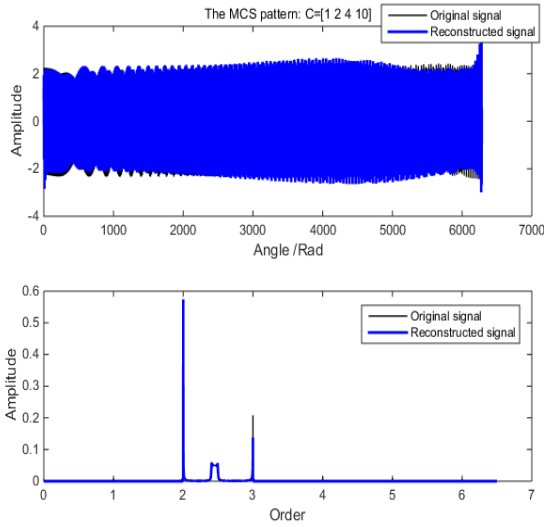
Fig.5. Reconstruction results

Furthermore, the angular-domain sampled signal is reconstructed based on the proposed method. The results in both angular domain and order domain are compared in Fig.5. We can see that good reconstructions are achieved.

B. Comparisons of Different Reconstruction Algorithms

Firstly, the sparsity of the simulated multi-band signal is equal to 3. Then the number of BTT probes (I) is chosen from 2 to L. For each I, 100 different MCS patterns (C) are randomly generated and the corresponding angular-domain BTT signals are sampled to calculate the averaged support reconstruction ratio (R). Here, the four reconstruction algorithms (OMP, MUSIC, BPDN and MFOCUSS) are applied respectively. In the end, the averaged support reconstruction ratio (R) is calculated under each algorithms and the results are shown as Fig.6. We can see that: i) Average of R increases with I for each algorithm, which testifies that increasing I can obviously improve the reconstruction performance; ii) When I is larger than 4, perfect reconstructions can almost be definitely carried out. On the other hand, when I is less than 4, the reconstruction errors will be large. This conclusion is consistent with the results of Table 1 and Eq.(26); iii) The averaged performances of MFOCUSS and BPDN are better than other two algorithms.
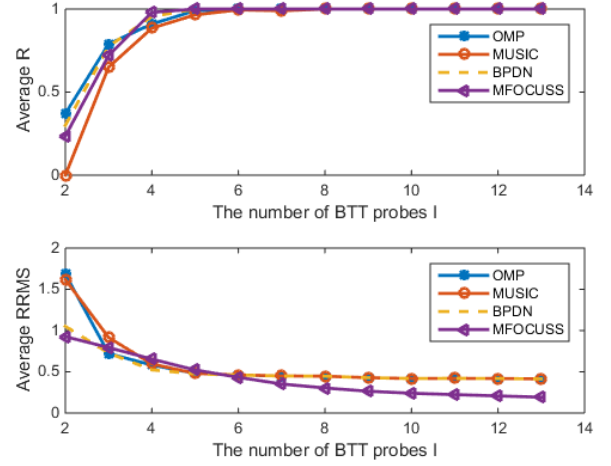
Fig. 6. Comparisons of four reconstruction algorithms.

C. Influences of Measurement Noises

In theoretical analysis, noises are not considered. In reality, measurement noises indeed exist, which will affect the reconstruction process. In order to investigate the influences, measurement noises with different SNRs are added into the simulated multi-band signal. 100 different MCS patterns (C) are randomly generated and the corresponding angular-domain BTT noised signals are sampled to calculate the averaged support reconstruction ratio (R). The results are shown in Fig.7. We can see that: i) Noises decrease the reconstruction performances of all the four algorithms; ii) When the SNR is low, the averaged performance of MFOCUSS is bad. Conversely, when the SNR is high, the averaged performance of MFOCUSS is the best; iii) Noises have the least influences on BPDN among the four algorithms.

In summary, the MUSIC method needs the maximum number of BTT probes for perfect reconstruction. The overall performance of BPDN is the best among the four algorithms according to the reconstruction error and sensitivity to noises. In addition, the most important is that BPDN can run without any prior information of the original signal, including the sparsity and the number of non-zero order slices.
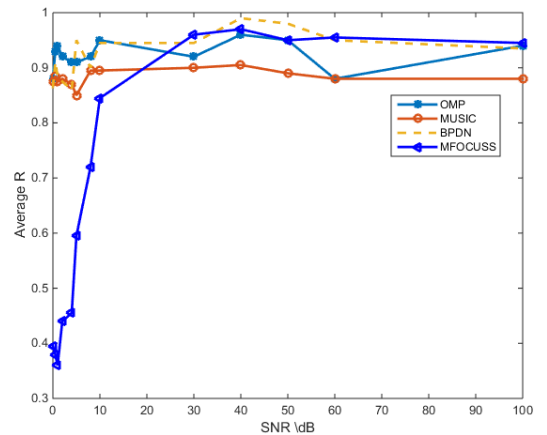
Fig. 7. Reconstruction results under different SNRs.

V. Conclusions

To deal with the problems of variable rotation speeds and under-sampled signals in the BTT method, this paper proposes an angular-domain CS reconstruction method in. The CS model of BTT measurement signals is derived in order domain. Two metrics are defined to characterize the reconstruction performance. By simulation, it is testified that the overall performance of BPDN is the best among the four algorithms according to the reconstruction error and sensitivity to noises. It must be noted that average rotating speeds in Eq.(4) may cause much bias under fast variable rotating speeds. Thus it is much valuable to investigate this problem in the future.

References

[1] Z. H. Song, Analysis of typical malfunction of aero engine, Beijing University of Aeronautics and Astronautics Press, 1993.

[2] A. V. Flotow, M. Mercadal, and P. Tappert, "Health monitoring and prognostics of blades and disks with blade tip sensors," Proc. IEEE Aerospace Conference, vol. 6, pp.433-446, 2000.

[3] A. Ghoshal, M. Sundaresan, M. Schulz, and P. Pai, "Structural health monitoring techniques for wind turbine blades," J. Wind Eng. Ind. Aerod., vol. 85, pp.309-324, 2000.

[4] Z. S. Chen, H. Sheng, Y. M. Xia, W. M. Wang, and J. He, "A comprehensive review on blade tip timing-based health monitoring: status and future," Mech. Syst. Signal Process., vol. 149, 107330, 2021.

[5] C. Lawson and P. Ivey, "Tubomachinery blade vibration amplitude measurement through tip timing with capacitance tip clearance probes," Sens. Actuators. A, vol.118, pp. 14-24, 2005.

[6] B. Salhi, J. Lardies, and M. Berthillier, "Identification of modal parameters and aeroelastic coefficients in bladed disk assemblies," Mech. Syst. Signal Process., vol. 23, pp. 1894–1908, 2009.

[7] Z. S. Chen, Y. M. Yang, B. Guo, and Z. Hu, "Blade damage prognosis based on kernel principal component analysis and grey model using subsampled tip-timing signals," Proc. Inst. Mech. Eng. Part C-J Mec., vol. 228, pp. 3178-3185, 2014.

[8] Z. S. Chen, Y. M. Yang, Y. Xie , B. Guo, and Z. Hu, "Non-contact crack detection of high-speed blades based on principal component analysis and Euclidian angles using optical-fiber sensors," Sensor Actuat. A-Phys., vol. 201, pp. 66-72, 2013.

[9] H. L. Xu, Z. S. Chen, Y. P. Xiong, Y. M. Yang, and L. M. Tao, "Nonlinear dynamic behaviors of rotated blades with small breathing cracks based on vibration power flow analysis," Shock Vib. 2016, 2016.

[10] H. L. Xu, Z. S. Chen, Y. M. Yang, L. M. Tao, and X. F. Chen, "Effects of crack on vibration characteristics of mistuned rotated blades," Shock Vib. 2017, 2017.

[11] J. Lin, Z. Hu, Z. S. Chen, Y. M. Yang, and H. L. Xu, "Sparse reconstruction of blade tip-timing signals for multi-mode blade vibration monitoring," Mech. Syst. Signal Process., vol. 85, pp. 250–258, 2016.

[12] Z. Hu, J. Lin, Z. S. Chen, Y. M. Yang, and X. J. Li, "A non-uniformly under-sampled blade tip-timing signal reconstruction method for blade vibration monitoring," Sensors, vol. 15, pp. 2419-2437, 2015.

[13] Y. Bresler, "Spectrum-blind sampling and compressive sensing for continuous-index signals," Proc. of Information Theory and Applications Workshop (ITA 2008), pp. 547-554, January 2008.

[14] R. Venkataramani, Y. Bresler, "Optimal sub-Nyquist nonuniform sampling and reconstruction for multiband signals," IEEE Trans. Signal Process., vol. 49, no. 10, pp. 2301-2313, 2001.

[15] D. Donoho and J. Tanner, "Precise undersampling theorems," Proc. of the IEEE, vol. 98, pp. 913 –924, June 2010.

[16] D. Donoho, "Compressed sensing," IEEE T. Inform. Theory, vol. 52, pp. 1289–1306, 2006.

[17] M. Rashidi, Non-uniform sampling and reconstruction of multi-band signals and its application in wideband spectrum sensing of cognitive radio, Chalmers University of Technology, Göteborg, Sweden, 2010.

[18] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," Magn. Rseon. Med., vol. 58, pp. 1182-1195, 2007.

[19] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressedive sampling," IEEE Signal Proc. Mag., vol. 25, pp. 83−91, 2008.

[20] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer, 2010.

[21] H. L. Xu, Z. S. Chen, Y. M. Yang, L. M. Tao, F. J. Guan, and H. F. Hu, "Damage detection in high-speed rotated blades by blade tip-timing method based on compressed sensing," Proc. of 2017 Prognostics and System Health Management Conference, IEEE, 2017.

[22] M. H. Pan, Y. M. Yang, F. J. Guan, H. F. Hu, and H. L. Xu, "Sparse representation based frequency detection and uncertainty reduction in blade tip timing measurement for multi-mode blade vibration monitoring," Sensors, vol. 17, pp. 1745, 2017.

[23] R. P. Spada and R. Nicoletti, "Applying compressed sensing to blade tip timing data: A parametric analysis," Proc. of the 10th International Conference on Rotor Dynamics (IFToMM 2018), vol. 3, pp. 121-134, 2018.

[24] Z. S. Chen, J. H. Liu, C. Zhan, J. He, and W. M. Wang, "Reconstructed order analysis-based vibration monitoring under variable rotation speed by using multiple blade tip-timing sensors," Sensors, vol. 18, pp.3235, 2018.

[25] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," IEEE Trans. Signal Process., vol. 53, pp. 2477~ 2488, 2005.

Online Intelligent Evaluation of Dispensing Quality Based on Entropy Weight Fuzzy Comprehensive Evaluation Method and Machine Learning

Liping Zhao*
State Key Laboratory for Manufacturing Systems Engineering Xi'an Jiaotong University
Xi'an, China
lipingzh@mail.xjtu.edu.cn

Xiangqian Cheng
State Key Laboratory for Manufacturing Systems Engineering Xi'an Jiaotong University
Xi'an, China
chengxiangqian@stu.xjtu.edu.cn

Yiyong Yao
School of Mechanical Engineering Xi'an Jiaotong University
Xi'an, China
yyyao@mail.xjtu.edu.cn

Abstract—**At present, the evaluation of dispensing quality is mainly manual. It has the shortcomings of strong subjectivity, low efficiency, low level of intelligence and inability to evaluate online. In this paper, an online intelligent evaluation method for dispensing quality based on entropy weight fuzzy comprehensive evaluation method combined with support vector machine(SVM) is proposed. In this method, the glue strip uniformity, glue strip width error, glue viscosity, etc. were determined as evaluation indexes by analyzing the glue performance, glue distribution state and glue strip width consistency. The visual detection method is used to detect the index value in real time. The entropy weight method is used to give appropriate weight to the indexes, and the fuzzy comprehensive evaluation method is used to infer the quality level of the dispensing process. The dispensing indexes-dispensing quality level mapping model based on support vector machine is established to realize the online intelligent evaluation of dispensing quality. Experimental results show that the method is effective.**

Keywords—dispensing, entropy weight, fuzzy comprehensive evaluation method, visual detection, support vector machine

I. INTRODUCTION

The digitization, networking and intelligence of the manufacturing industry are the core technologies of the new round of industrial revolution, and should serve as the key point and main direction of "Made in China 2025" [1]. Dispensing technology is a key technology in microelectronic packaging, it can be used to form dots, lines, surfaces and various graphics, and is widely used in die attachment, encapsulation and coating [2]. Some papers have studied the intellectualization of dispensing process.

Peng determines whether there is broken glue or wavy glue by processing the glue image [3]. Tan proposed a quality detection method of LED surface defects in dispensing back-end based on image preprocessing, image matching [4]. Zhang proposed an calibration method for dispensing amount. the support vector machine model of the actual dispensing amount is established by using the dispensing speed and the set value as input vectors, then the revised set value and calibration coefficient are calculated by iterative method [5]. Liu proposed a compound fuzzy control strategy which can adjust valve open-time in each dispensing cycle. It makes the system accurately dispense smaller volume of liquid and immune to liquid viscosity, pressure fluctuation and some other disturbances [6]. Monroe Kennedy used visual feedback to detect the fluid height in the beaker, and the design of the controller enables the robot arm to achieve precise dispensing of liquids [7]. Cheng designed a positioning system that uses machine vision instead of manual input to automatically measure the dispensing position [8]. Zhang designed a visual dispensing system that can automatically identify the parts [9]. Cha used a convolutional neural network to detect defects such as wider glue and broken glue [10].

It can be found that the current intelligent dispensing process mainly relies on machine vision technology, machine learning and deep learning algorithm, and there are few studies on the evaluation of dispensing quality. In this paper, entropy weight fuzzy comprehensive evaluation method, visual detection method, and support vector machine (SVM) are combined to achieve online intelligent evaluation of dispensing quality.

II. ANALYSIS OF DISPENSING QUALITY EVALUATION INDEX

The research object of this paper is glue strip. Dispensing quality is affected by many factors such as dispensing speed, viscosity, temperature, gas pressure and so on. There may be defects such as wider glue, narrower glue, broken glue, satellite drops, bubbles and glue width errors. Focusing on the factors affecting the quality of dispensing and common quality problems, the evaluation index system for dispensing quality is established, as shown in Fig.1.
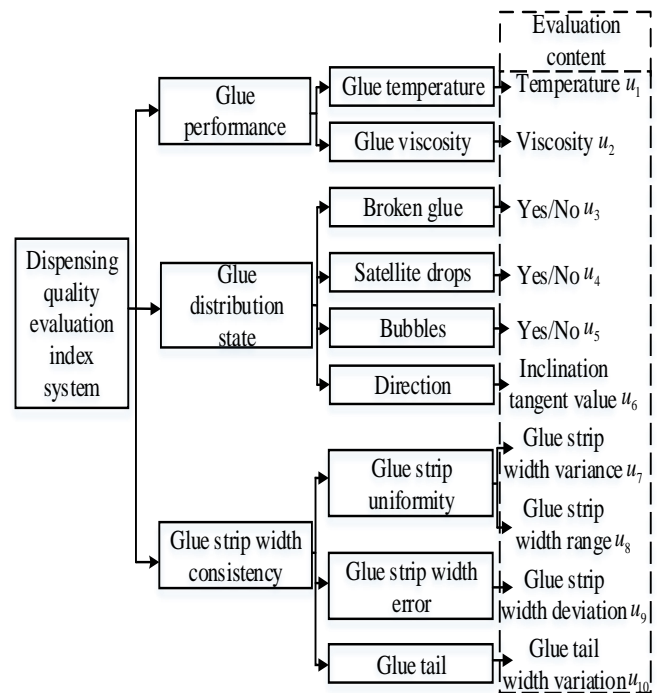
Fig.1. Dispensing quality evaluation index system

---

*Liping ZHAO is the corresponding author. (e-mail: lipingzh@mail.xjtu.edu.cn)

The glue performance mainly refers to the viscosity properties of the glue itself and temperature. Generally speaking, the viscosity decreases with increasing temperature. Too low viscosity affects the firmness of the bonding. If the viscosity is too high, the glue will easily remain in the needle tube, which will affect the accuracy of the next dispensing. The evaluation of viscosity and temperature can ensure the quality of the glue.

The glue distribution state is defined as the state of the glue itself and the glue position distribution along the length of the glue strip on substrates. The ideal glue distribution state is bubble-free, continuous and evenly distributed, but in the actual dispensing process, phenomena such as broken glue, satellite drops, and bubbles often occur. At the same time, the directional error of glue distribution may lead to the change of adhesive position and the decrease of adhesive firmness. The rationality of the glue distribution can be ensured by evaluating the presence or absence of the above phenomena and the directionality of the glue distribution. In the evaluation content, the inclination tangent value refers to the tangent value of the acute angle which is formed by the glue strip and its ideal position.

For one glue strip, the width consistency is defined as the glue distribution state along the width direction of glue strip. The ideal width of one glue strip has a high width consistency and has a small error from the theoretical width. In the actual dispensing process, due to the influence of factors such as dispensing speed and air pressure, there will often be poor consistency of the glue strip width, large deviation from the theoretical glue width, and glue tail. Statistical analysis and evaluation of the glue width are used to ensure the consistency of the glue strip width. In the evaluation content, the glue strip width deviation is defined as the difference between the theoretical width of glue strip and the mean width obtained by measurement, the glue tail width variation is defined as the difference between the end width of glue strip and the width of glue strip at L cm from the end. The value of L is approximately equal to one-sixth of the strip's length.

III. CONSTRUCTION OF DISPENSING QUALITY EVALUATION MODEL

If satellite drops, bubbles and broken glue occur during the dispensing process, the dispensing quality is considered to be unqualified. Otherwise, the dispensing quality level is evaluated according to the dispensing quality evaluation model proposed in this paper. The evaluation process of dispensing quality is shown in Fig.2.

Because there are many evaluation indexes of glue strip quality and the action intensity of each evaluation index is different, it's difficult to quantitatively describe the dispensing quality. As an objective weighting method, the entropy weight method can determine the objective weight of the index according to the variability of each evaluation index, the fuzzy comprehensive evaluation method is a comprehensive evaluation method based on fuzzy mathematics, which can quantify objects that are difficult to quantify [11]. Therefore, the entropy weight method and fuzzy comprehensive evaluation method are combined to evaluate the dispensing quality.
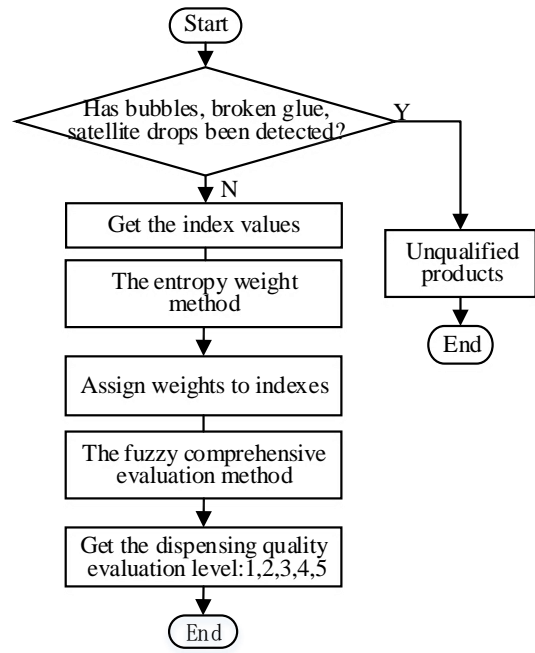
Fig.2 Dispensing quality evaluation process

A. Determine the Weight of Evaluation Indexes

Step 1. Construct and standardize the evaluation matrix. The evaluation matrix of the evaluation model is as follows:

$$P' = (p_{ij}')_{m \times n} = \begin{bmatrix} p_{11}' & p_{12}' & \cdots & p_{1n}' \\ p_{21}' & p_{22}' & \cdots & p_{2n}' \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1}' & p_{m2}' & \cdots & p_{mn}' \end{bmatrix} \quad (1)$$

Where $p_{ij}' (i = 1, 2, \cdots m, j = 1, 2, \cdots n)$ is the value of the index, i is the serial number of the glue strip to be evaluated, j is the serial number of the evaluation index, m and n are the number of glue strips and the number of evaluation indexes respectively. After the evaluation matrix is established, it needs to be standardized by some dimensionless processing. If the larger the value of a certain index, the better the quality of the glue strip, then the index is a benefit index, otherwise, the index is a cost index. The standardized methods of the benefit index and the cost index are as follows:

$$p_{ij} = (p_{ij}' - \min p_{ij}') / (\max p_{ij}' - \min p_{ij}') \quad (2)$$

$$p_{ij} = (\max p_{ij}' - p_{ij}') / (\max p_{ij}' - \min p_{ij}') \quad (3)$$

Where p_{ij} is the standardized index value, $\max p_{ij}'$ and $\min p_{ij}'$ are the maximum index value and the minimum index value respectively. P is the standardized matrix.

$$P = (p_{ij})_{m \times n} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{bmatrix} \quad (4)$$

Step 2. Determine the weight of each index. According to the definition of entropy value and entropy weight, the entropy value and entropy weight of the index j are as follows:

$$H_j = -\frac{1}{\ln m} \sum_{i=1}^{m} \frac{p_{ij}}{(\sum_{i=1}^{m} p_{ij})} \ln \frac{p_{ij}}{(\sum_{i=1}^{m} p_{ij})} \tag{5}$$

$$w_j = \frac{1 - H_j}{n - \sum_{j=1}^{n} H_j} \tag{6}$$

In (5), $0 \le H_j \le 1$. Assumption when $p_{ij} / (\sum_{i=1}^{m} p_{ij}) = 0$,

$p_{ij} / (\sum_{i=1}^{m} p_{ij}) \ln p_{ij} / (\sum_{i=1}^{m} p_{ij}) = 0$. In (6), $0 \le w_j \le 1$, $\sum_{j=1}^{n} w_j = 1$.

B. Fuzzy evaluation of glue strip quality

Step 1. Determine the decision set. The index set is a set of n evaluation indexes. The index set of the glue strip i is: $P_i = \{p_{i1}, p_{i2} \cdots p_{in}\}$. The decision set refers to the set of all possible evaluation results of each glue strip to be evaluated. In this paper, the decision set is set to five levels, that is, $V = \{v_1, v_2, v_3, v_4, v_5\} = \{1, 2, 3, 4, 5\}$.

Step 2. Construct fuzzy relation matrix. The fuzzy matrix can be calculated by the membership function. In this paper, the triangular distribution membership function is selected. The membership functions corresponding to the five levels in the decision set are as follows:

$$r_{v1}(p_{ij}) = \begin{cases} \dfrac{p_{ij} - 0.4}{0.6} & 0.4 \le p_{ij} \le \mu_1 \\ 0 & others \end{cases} \tag{7}$$

$$r_{v2}(p_{ij}) = \begin{cases} \dfrac{p_{ij} - 0.15}{0.6} & 0.15 \le p_{ij} \le \mu_2 \\ \dfrac{1.35 - p_{ij}}{0.6} & \mu_2 \le p_{ij} \le 1 \\ 0 & others \end{cases} \tag{8}$$

$$r_{v3}(p_{ij}) = \begin{cases} \dfrac{p_{ij} + 0.1}{0.6} & 0 \le p_{ij} \le \mu_3 \\ \dfrac{1 - p_{ij}}{0.6} & \mu_3 \le p_{ij} \le 1 \end{cases} \tag{9}$$

$$r_{v4}(p_{ij}) = \begin{cases} \dfrac{p_{ij} + 0.35}{0.6} & 0 \le p_{ij} \le \mu_4 \\ \dfrac{0.85 - p_{ij}}{0.6} & \mu_4 \le p_{ij} \le 0.85 \\ 0 & others \end{cases} \tag{10}$$

$$r_{v5}(p_{ij}) = \begin{cases} \dfrac{0.6 - p_{ij}}{0.6} & \mu_5 \le p_{ij} \le 0.6 \\ 0 & others \end{cases} \tag{11}$$

Among them, μ_1, μ_2, μ_3, μ_4, and μ_5 are the expected distribution values corresponding to the five levels in the decision set. Here, the values of μ_1, μ_2, μ_3, μ_4, and μ_5 are set to 1, 0.75, 0.5, 0.25, and 0, respectively. According to the above 5 membership functions, the membership degree of the index can be obtained and are expressed as $[r_{j1}, r_{j2}, r_{j3}, r_{j4}, r_{j5}]$. Therefore, the fuzzy evaluation matrix of the glue strip i can be expressed as:

$$R_i = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{15} \\ r_{21} & r_{22} & \cdots & r_{25} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{n5} \end{bmatrix} \tag{12}$$

Step 3. Calculate the comprehensive evaluation vector. The comprehensive evaluation vector of the glue strip is related to the fuzzy evaluation matrix and the index weight vector, indicating the degree of membership of the glue strip to the 5 levels, which can be expressed as:

$$B_i = W \circ R_i = [b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5}] \tag{13}$$

Where $\circ$ represents a fuzzy composition operator. According to the operator, the formula for calculating each element in the comprehensive evaluation vector is as follows:

$$b_{ik} = \sum_{j=1}^{n} w_j r_{jk} \ (k = 1, 2, 3, 4, 5) \tag{14}$$

Step 4. Evaluate the glue strip. According to the principle of maximum membership to evaluate the quality of glue strips.

IV. ONLINE INTELLIGENT EVALUATION OF DISPENSING QUALITY

To realize the online evaluation of dispensing quality, the first step is to realize the real-time detection of dispensing index data. In this paper, the real-time detection of some dispensing index data is realized through visual detection technology. The intelligent evaluation of dispensing quality is a process in which the real-time dispensing index data is analyzed automatically and the dispensing quality levels are inferred based on past experience, in this paper, the dispensing indexes-dispensing quality evaluation level mapping model based on SVM is established to realize the process.

A. Real-time detection of dispensing index data

1) Real-time detection of glue strip width

The glue strip image can be obtained by industrial camera and its edge image is obtained by graying, filtering, image segmentation and edge extraction, as shown in Fig.3. By traversing the edge image line by line, the pixel coordinates of the two edges of the glue strip can be obtained, expressed as (x_{pl}, y_{pl}) and (x_{pr}, y_{pr}). Where $p(p = 1, 2 \cdots h)$ is the line number, h is the height of the glue strip, l refers to the left edge, r refers to the right edge. Then $b = x_{pl} - x_{pr}$ is the width of the glue strip corresponding to each row. The actual distance represented by a single pixel, that is, the calibration coefficient C, can be obtained by processing the ruler image. Therefore, the glue strip width can be calculated by the formula $B = Cb$.

(a) Original glue strip image   (b) The edge image of the glue strip.

Fig.3. Glue strip image

2) *Real-time detection of the tangent value of the inclination*

The top and bottom ends of the left edge can be obtained through the edge image of the glue strip, that is, $(x_{0l}, y_{0l}), (x_{(h-1)l}, y_{(h-1)l})$ .Then the tangent value of the inclination can be calculated by the formula:

$$t_l = | x_{0l} - x_{(h-1)l} | / h \tag{15}$$

Similarly, the right one can be calculated by the formula:

$$t_r = | x_{0r} - x_{(h-1)r} | / h \tag{16}$$

The larger one between t_l and t_r is selected as the final value.

The contour of the glue strip to be tested can be obtained By means of image de-noising, template analysis, image segmentation and morphological processing, the difference set with the qualified glue's centerline was operated to check for the broken glue, bubbles and tell if satellite drops existed according to the width of glue contour [3].

B. Online Intelligent Evaluation of Dispensing Quality

The quality level of the glue strip can be obtained by the entropy weight fuzzy comprehensive evaluation method. With the evaluation index value of the glue strip as the input and the corresponding quality level as the output to train the SVM, the intelligent evaluation model of dispensing quality can be established. Since there are 5 levels of the glue strip quality, based on the principle of one-versus-rest, five binary sub-classifiers are established, and then combined into a multi-class classifier, as shown in Fig.4.
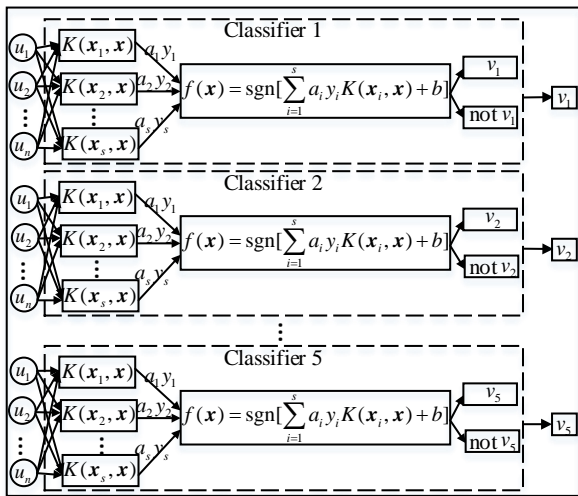
Fig.4. Intelligent evaluation model

In Fig.4, $K(x_i, x)$ is a kernel function. In this paper, the radial basis function is selected as the kernel function, and its function expression is:

$$K(x_i, x) = \exp(-\frac{\| x_i - x \|^2}{\sigma^2}) \tag{17}$$

Where x_1 , $x_2 \cdots$, x_s are support vectors, $x = (u_1, u_2, \cdots u_n)$ is an input vector composed of evaluation index values. $a_1 y_1$, $a_2 y_2 \cdots a_s y_s$ are weights. $f(x)$ is a linear combination of several middle layer nodes, a_i is the Lagrange multiplier, y_i is the classification label, b is the classification threshold [12]. Through visual detection, the evaluation index data are input into the intelligent evaluation model of dispensing quality in real time, and the corresponding evaluation levels of the glue strips are output. In this way, the online intelligent evaluation of the dispensing quality is realized. At the same time, the input and output are added to the training set to continuously train the intelligent evaluation model.

V. CASE STUDY

In this paper, the evaluation of dispensing quality was carried out with the focus on the consistency of glue strip width and the directionality of glue distribution. Through visual detection, the index data of 30 glue strips are obtained, part of the data are shown in TABLE Ⅰ.

TABLE I.   EVALUATION INDEX VALUES OF GLUE STRIPS

Glue Number	Evaluation Indexs				
	u_6	u_7	u_8 (mm)	u_9 (mm)	u_{10} (mm)
1	0.026	0.074	1.786	1.722	0.571
2	0.022	0.611	2.929	0.244	0.214
3	0.005	0.458	3.000	0.817	2.571
4	0.060	0.135	2.429	0.912	0.286
5	0.070	2.530	6.786	0.159	0.571
6	0.057	0.277	2.143	0.492	1.929
7	0.002	0.083	1.214	0.752	0.000
8	0.029	0.240	4.500	1.550	0.857
9	0.010	0.560	2.643	2.185	0.714
10	0.026	0.050	1.071	1.883	0.571

The data are standardized by formula (3). According to formula (5) and (6), the entropy weight of each index is as follows:

$$W^T = [0.158\ 0.086\ 0.211\ 0.284\ 0.261]$$

Taking the glue strip 1 as an example, the fuzzy evaluation matrix can be obtained by formula (7)~(12):

$$R_1 = \begin{bmatrix} 0.402 & 0.819 & 0.598 & 0.348 & 0.000 \\ 0.972 & 0.612 & 0.029 & 0.000 & 0.000 \\ 0.792 & 0.792 & 0.208 & 0.000 & 0.000 \\ 0.030 & 0.447 & 0.864 & 0.720 & 0.303 \\ 0.630 & 0.954 & 0.370 & 0.120 & 0.000 \end{bmatrix}$$

According to formula (13) and (14), the comprehensive evaluation vector of glue strip 1 is as follows:

$$B_1 = [0.488\ 0.725\ 0.483\ 0.290\ 0.086]$$

According to the principle of maximum membership, the evaluation result of the glue strip 1 is level 2. Similarly, the evaluation results of the remaining glue strips can be obtained. Some glue strips and their evaluation results are shown in Fig.5.
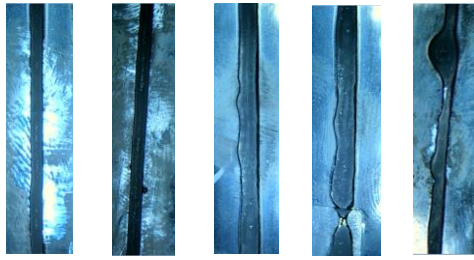
(a) Level 1 (b) Level 2 (c) Level 3 (d) Level 4 (e) Level 5

Fig.5. Evaluation results

It can be seen from Fig.5 that after comprehensive consideration of multiple indexes, the evaluation results are not only objective but in line with human subjective evaluation, and the evaluation results are acceptable.

In Python environment, the intelligent evaluation model is simulated. The experimental data of the first 20 glue strips are used as the training set to train the SVM and the experimental data of the last 10 glue strips are used as the test set to test the generalization ability of the intelligent evaluation model. The experiment results are shown in TABLE Ⅱ.

TABLE II.   PREDICTION RESULTS OF TEST SAMPLES

Glue Number	Actual Level	Prediction Level
21	2	2
22	2	2
23	2	2
24	2	2
25	2	2
26	2	2
27	2	2
28	1	2
29	2	2
30	2	2

It can be seen from TABLE Ⅱ that in the case of fewer training samples, the prediction results of the intelligent evaluation model for the test samples are accurate, the quality of the glue strips can be effectively evaluated according to its index value.

VI.   CONCLUSION

Focusing on the factors that affect the quality of dispensing and the common problems of dispensing quality, a dispensing quality evaluation index system was constructed from three aspects: glue performance, glue distribution state and glue strip width consistency. The entropy weight fuzzy comprehensive evaluation method was used to evaluate the quality of dispensing, the experimental results showed that the evaluation results were not only objective but in line with human subjective evaluation. The visual detection method was used to detect the index data, and an intelligent evaluation model for dispensing quality based on SVM was constructed. During the dispensing process, the index data are input into the intelligent evaluation model in real time through visual detection, and the corresponding evaluation results are output, realizing the online intelligent evaluation of the dispensing quality.

REFERENCES

[1]   J. Zhou, "Intelligent Manufacturing—main direction of "Made in China 2025"," China Mechanical Engineering, vol. 26, no. 17, pp. 5-16, 2015.

[2]   X. Li, H. Wang, " Classification and development trend of fluid dispensing technology," CHENGSHI JIANSHE YU SHANGYE WANGDIAN, no. 3, pp. 26-28, 2009.

[3]   Y. Peng, Y. Liu and D. Zhang, " Design of Glue Dispensing Quality Detection System Based on HALCON," PACKAGING ENGINEERING, vol. 39, no. 15, pp. 204-209, 2018.

[4]   J. Tan, L. Li, Y. Wang, F. Mo, J. Chen, L.Zhao and Y. Xu, "The Quality Detection of Surface Defect in Dispensing Dack-End Based on HALCON," 2016 International Conference on Cybernetics, Robotics and Control (CRC), Hong Kong, 2016, pp. 95-98.

[5]   J. Zhang, J. Hu, Y. Cao and H. Chi, "Study on a method of dispensing calibration for photoresist pump based on support vector machine," 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), Shenyang, 2015, pp. 1552-1556.

[6]   Y. Liu, Y. Yao, L. Chen and L. Sun, "A closed-loop intelligent control strategy for precise non-contact liquid dispensing," 2010 IEEE International Conference on Mechatronics and Automation, Xi'an, 2010, pp. 957-962.

[7]   M. Kennedy, K. Queen, D. Thakur, K. Daniilidis and V. Kumar, "Precise dispensing of liquids using visual feedback," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, 2017, pp. 1260-1266.

[8]   F. Chrng, X. Zhang and J. Zhang, "A Positioning System for Dispensing Machine Based on Machine-Vision," Machinery Design & Manufacture, no. 3, pp. 101-104, 2013.

[9]   K. Zhang, H. Wang, X. Chen, N. Cai, Y. Zeng and G. He, " Visual Dispensing System Based on Automatic Recognition of Workpieces ," Modular Machine Tool & Automatic Manufacturing Technique, no. 7, pp. 43-47, 2018.

[10]   G. Cha, G. Hu, " Detection of dispensing defects based on deep learning," Electronic Technology & Software Engineering, no. 13, pp. 49-52, 2019.

[11]   W. Ouyang, Q. Cheng, Y. Jin, Z. Liu, X. Liu and B. Wang, " Performance Evaluation of Water Lubricated Stern Bearing Based on Entropy Weight Fuzzy Comprehensive Evaluation Method," China Mechanical Engineering, in press.

[12]   G. Wang, Q. Li, X. Qin, X. Yu, Y. Cui and D. Peng, "Application of Support-Vector-Machine in Tool Wear of Multi-Stage Monitoring," JOURNAL OF TIANJIN UNIVERSITY, vol. 44, no. 1, pp. 35-39, 2011.

A Deep Reinforcement Learning Based Control Approach for Suspension Systems of Maglev Trains

Yougang Sun *
National Maglev Transportation Engineering R&D Center, and Institute of Rail Transit
Tongji University
Shanghai, China
1989yoga@tongji.edu.cn

Junqi Xu
National Maglev Transportation Engineering R&D Center
Tongji University
Shanghai, China
xujunqi@tongji.edu.cn

Chen Chen
National Maglev Transportation Engineering R&D Center
Tongji University
Shanghai, China
c_chen_brightness@tongji.edu.cn

Wei Hu
CRRC Zhuzhou locomotive Co., Ltd
Zhuzhou Hunan, China
1094065783@qq.com

Abstract—The magnetic suspension control system is one of the core components of the maglev trains. However, because of the system's unstable open loop, strong nonlinearity, and model uncertainty, the design of maglev suspension control method is challenging. In this paper, a third-order maglev train suspension system dynamics model is established firstly. Then the affine nonlinear model of the magnetic suspension system through the nonlinear coordinate transformation theorem is obtained. Subsequently, without making any linear approximation, the nonlinear integral sliding mode controller (NISMC) is directly developed and the stability analysis is performed. To eliminate the influence of system disturbance on control performance, RBF neural network and Actor-Critic algorithm combined to construct a modified deep reinforcement learning method, which is used to optimize controller parameters in real time and enhance system robustness. Numerical simulation results are provided to demonstrate the effeteness of the proposed deep reinforcement learning method.

Keywords—*Deep reinforcement learning, maglev trains, sliding model control, suspension systems*

I. INTRODUCTION

With the gradual improvement of the public transportation, people's expectations for more environmental protection and green rail transit are further stimulated. The expectation of 600 km / h or even higher travel speed, lower operation and maintenance cost, lower pollution, and other aspects provide opportunities for the development of maglev rail transit. For maglev, a new type of rail transit, due to its non-adhesive operation characteristics, it overcomes the mechanical wear and friction between the rail and the wheels, and can obtain a higher speed in theory. In 2003, The first commercial maglev line between Longyang Road and Pudong International Airport was opened in Shanghai[1]. The vehicle used in this line is TR type maglev vehicle developed by MBB. This type of maglev vehicle adopts the principle of electromagnetic suspension and maintains a rated small clearance of about 8-10 mm [2]. TR maglev is famous for its longest commercial operation time and the most mature technology in the world.

At present, more and more researchers have discussed and analyzed the stability of suspension from the perspective of advanced control algorithms. Wang et al. [3] utilized a gain control table to adjusted the control gains online adaptively, which can improve the system robustness when the track stiffness on the flexible track is insufficient. However, because the model of the controller is designed on the linear

approximation model, although it has been improved compared with the linear model, when it is far from the balance point, it will also face the phenomenon of low control performance. Chen et al [4] proposed a Takagi-Sugeno fuzzy model to deal with parameter uncertainty and external disturbance. On this basis, an adaptive fuzzy control approach is utilized for suspension control with a particle swarm optimization algorithm. In order to improve the dynamic properties of the passive suspension, Wang et al [5] proposed a suspension system based on semi-active inertia. A relative acceleration-velocity controller is developed based on the inertial elements mechanical characteristics and relative acceleration-relative velocity of the system. The simulation results demonstrate that the semi-active inertial suspension system with the relative acceleration relative velocity control strategy has better dynamic performance under the dynamic load. Xu et al. [6] studied the dynamic characteristics of the maglev vehicle under random irregularity based on a flexible track. Based on the decomposition of track force into a segmented chain structure, a method of vertical suspension stability analysis of maglev vehicle was proposed. The discrete form of the track segment chain structure and the motion equations of track structure are established. The random excitation generated by track irregularity is transformed into system input excitation by virtual excitation method, and the coupling mechanism of feedback control gains on the maglev system stability is analyzed. Nebiyeleul et al. [7] proposed a control scheme based on proportional integral observer (PIO), which still uses simple conventional cascade controller to enhance the robustness. It is verified that the controller can maintain the normal performance even in this bad condition. In addition, the compensation performance of the observer is analyzed based on the singular perturbation theory. Considering the track elasticity, the model of the suspension system will be more complex, and the control algorithm is difficult to be realized in engineering. For this reason. Li and others [9] proposed a structure of electromagnetic levitation (EMS) Maglev Train Based on V-shaped track, which uses a set of on-board magnets and long ground stator coils to achieve the function of levitation guidance and traction. According to the characteristics of independent suspension motion and coupling of guidance and rolling motion, the control strategy of independent design of suspension controller, guidance and rolling controller is proposed. The simulation results show that the suspension, guidance and rolling motion can achieve stable control. To research the influence of time-delays of EMS maglev

---

* Yougang Sun is the corresponding author. (e-mail: 1989yoga@tongji.edu.cn).

suspension controller on Maglev stability and solve the problem of maglev suspension instability, Chen et al. [10] took the time-delay in the closed-system as a parameter, and judged the characteristic root distribution of the system by Routh Hurwitz stability criterion. In addition, it gave the method of calculate the critical value for Hopf bifurcation. When the time-delay in the closed-system is greater than the calculated value, the system will be unstable.

However, in the suspension control process of maglev trains, once the controller parameters are determined, online adjustment cannot be achieved. In the field of artificial intelligence (AI), reinforcement learning can be seen everywhere since it is both a method of behavior optimization and a computing tool [11-12]. Among them, the Actor-Critic (AC) learning [13-14], as a typical reinforcement learning algorithm, can be utilized for strategy approximation of continuous action space or continuous state. This article takes advantage of the model-free online learning of RBF neural network and AC learning, and proposes a nonlinear integral sliding mode controller (SMC) based on the improved deep learning approach. This approach not only has the characteristics of fast convergence speed, but also can control the parameter adjustment in real time. The simulation results demonstrate that the proposed control law can ensure the airgap of the maglev train system track the given signal quickly and accurately, and remove the chattering phenomenon in the controller.

The article is organized as follows. In Section II, the dynamic model and control objective is given. Section III designs a deep reinforcement learning based controller. We include the simulation results in Section IV Section V provides the conclusions and outlooks of this paper.

II. PROBLEM STATEMENT

As illustrated in Fig. 1, the state space expression of the EMS maglev vehicle system with disturbance can be derived as below [15-16]:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = -\dfrac{\mu_0 A_m N_m^2}{4m}\left[\dfrac{x_3(t)}{x_1(t)}\right]^2 + g + f_d \\ \dot{x}_3(t) = \dfrac{x_2(t)x_3(t)}{x_1(t)} + \dfrac{2x_1(t)}{\mu_0 N_m^2 A_m}(u_m(t) - x_3(t)R_m) \\ y(t) = x_1(t) \end{cases} \quad (1)$$

where, the system status is $x_1(t) = x_m(t)$, $x_2(t) = \dot{x}_m(t)$ and $x_3(t) = i_m(t)$. $x_m(t)$ is the air gap between the electromagnet and the track; A_m is the effective magnetic pole area; μ_0 is the air permeability; R_m is the electromagnet coil resistance; the electromagnet coil current and electromagnet coil voltage are given as $i_m(t)$ and $u_m(t)$.
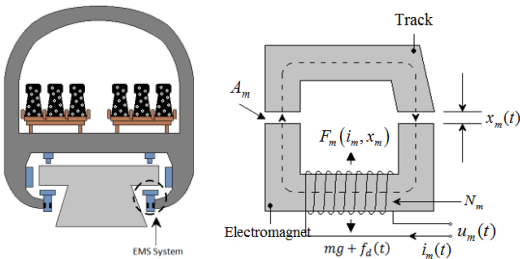
Fig. 1. Maglev train suspension system

A. Coordinate Change.

The dynamic model of the maglev system is nonlinear, which is not convenient for subsequent controller design. We can utilize coordinate transformation method to transform the nonlinear model to affine model of the magnetic suspension system. The new transformed coordinates are selected as state variables:

$\boldsymbol{\eta} = \begin{bmatrix} \eta_1 & \eta_2 & \eta_3 \end{bmatrix}^T \in \Omega_\eta$, $\Omega_\eta \in R^3$ is a compact set of R^3.

where:

$$\begin{cases} \eta_1 = x_1 \\ \eta_2 = x_2 \\ \eta_3 = g - \dfrac{\mu_0 A_m N_m^2}{4m}\left(\dfrac{x_3}{x_1}\right)^2 \end{cases} \quad (2)$$

Further, it is not difficult to get:

$$\begin{cases} \dot{\eta}_1 = \eta_2 \\ \dot{\eta}_2 = \eta_3 \\ \dot{\eta}_3 = -\dfrac{\mu_0 A_m N_m^2}{2m}\left(\dfrac{x_3 \dot{x}_3}{x_1^2}\right) + \dfrac{\mu_0 A_m N_m^2}{2m}\left(\dfrac{x_3^2 \dot{x}_1}{x_1^2}\right) \end{cases} \quad (3)$$

In order to facilitate the controller design and analysis, \ k and L are introduced as below:

$$k = \frac{\mu_0 A_m N_m^2}{4}, \quad L = \frac{2k}{x_1}$$

According to equation (3), $\dot{\eta}_3(t)$ can be rewritten as below:

$$\dot{\eta}_3 = \frac{2k}{m}\left(\left(1 - \frac{2k}{Lx_1}\right)\frac{x_2 x_3^2}{x_1^3} + \frac{R}{L}\frac{x_3^2}{x_1^2}\right) - \frac{2kx_3}{Lmx_1^2}u_m \quad (4)$$

Therefore, considering the disturbance and parameter perturbation, the affine model of the magnetic suspension system in the new coordinate system can be obtained as follows:

$$\begin{cases} \dot{\eta}_1 = \eta_2 \\ \dot{\eta}_2 = \eta_3 \\ \dot{\eta}_3 = f(\eta) + g(\eta)u_m + f_d \\ y = \eta_1 \end{cases} \quad (5)$$

where, $f(\boldsymbol{\eta}) = \dfrac{2k}{m}\left(\left(1 - \dfrac{2k}{Lx_1}\right)\dfrac{x_2 x_3^2}{x_1^3} + \dfrac{R}{L}\dfrac{x_3^2}{x_1^2}\right)$, $g(\boldsymbol{\eta}) = -\dfrac{2kx_3}{Lmx_1^2}$.

B. Control Objective.

The control goal of this section is that considering the magnetic suspension dynamic model (1), under the existence of external disturbance and parameter uncertainties, an active control strategy is designed to achieve the suspension tracking control of the maglev train. That is to say, the state $x_m(t)$ of the system can track the reference position command signal $r(t)$. And it can make the tracking error of $|x_m(t) - r(t)|$ converge to the neighborhood of the zero point. Before designing the controller, the following assumptions need to be made:

Assumption 1 The state of the system (1) is measurable, and due to the physical properties of the maglev train, the output state quantity has a limited upper limit value, that is

$\{x_m, \dot{x}_m, i_m || x_m| < \bar{\delta}_1, |\dot{x}_m| < \bar{\delta}_2, |i_m| < \bar{\delta}_3\}$, where $\bar{\delta}_i$ is an unknown positive number, $i = 1, 2, 3$.

Assumption 2 The expectation command $r(t)$ and its first and second derivatives are limited, that is $\{max(|r|, |\dot{r}|, |\ddot{r}|) \leq \bar{\delta}_r\}$, where $\bar{\delta}$ is a unknown positive value.

Assumption 3 External disturbance f_d is bounded, that is $|f_d| \leq f_{d \max}$, where $f_{d \max}$ is a unknown positive value. The external disturbances (such as strong winds) cannot be infinite, so there will always be an upper bound of θ.

III. DEEP REINFORCEMENT LEARNING BASED CONTROLLER DESIGN AND ANALYSIS

A. Design of Nonlinear Integral SMC

The EMS maglev vehicle system states are chosen as below:

$$\eta_1 = x_1 = x_m, \eta_2 = x_2 = \dot{x}_m(t)$$

System error and error change rate can be obtained as:

$$\begin{cases} e = \eta_1 - r \\ \dot{e} = \dot{\eta}_1 - \dot{r} = \eta_2 \end{cases} \quad (6)$$

where, r is a reference trajectory required to be tracked by output y, that is, the target suspended air gap, and the system error $e = \eta_1 - r$ converges to zero.

Based on the equations of (6), the integral sliding mode surface is developed as follows:

$$S \triangleq \left\{ (e, \dot{e}, \ddot{e}) | c_1 e + c_2 \dot{e} + \ddot{e} = c_0 \int_0^t e dt \right\} \quad (7)$$

where, c_1, c_2, $c_0 \in R^+$ is the positive control gain. Substituting (2), (5), and (6) into (7), the specific expression of the sliding mode surface is:

$$S(\boldsymbol{\eta}, t) = c_1 e(t) + c_2 \eta_2(t) + \eta_3(t) + c_0 \int_0^t e(t) dt \quad (8)$$

It describes the distance between the sliding mode surface and the system states. It can be learned from equation (8) that when the system state is on the sliding mode surface, which is $S(\boldsymbol{\eta}, t) = 0$, $e = \eta_1(t) - r$ can be guaranteed to converge to 0.

Derivation of equation (6) and substituting into (2), (5) gives:

$$\dot{S}(\eta, t) = c_1 \eta_2(t) + c_2 \eta_3(t) + f(\eta) + g(\eta) u_m(t) + c_0 e(t) \quad (9)$$

Selecting the exponential approach rate to design the controller, and it is $\dot{S}(\eta, t) = -(W + \theta) sgn(S) - \kappa S$, can obtain the sliding mode variable structure control law as:

$$u_m(\boldsymbol{\eta}, t) = -[g(\boldsymbol{\eta})]^{-1} \begin{bmatrix} f(\boldsymbol{\eta}) + c_1 \eta_2 + c_2 \eta_3 + c_0 e + (W + \theta) \mathrm{sgn}(S) \\ + \kappa \left(c_1 e + c_2 \eta_2 + \eta_3 + c_0 \int_0^t e(t) dt \right) \end{bmatrix} \quad (10)$$

where, W and $\kappa \in R^+$ represent the constant arrival coefficient and exponential arrival coefficient.

The Lyapunov function is defined as follows:

$$V(x) = \frac{1}{2} S^2 \quad (11)$$

This function is positive semidefinite, and the two sides of the formula can be derived:

$$\dot{V}(x) = S \cdot \dot{S} = -S \cdot (W + \theta) sgn(S) \quad (12)$$

$$= -(W + \theta)|S| - \kappa \cdot S^2 \leq 0$$

According to Lyapunov technology, the system is stable in the sense of Lyapunov.

B. Control Parameter Optimization based on Reinforcement Learning

According to the stability analysis process, the satisfaction $|f_d| \leq \theta$ of disturbance $f_d(x, t)$ is an important condition for Lyapunov stability. As the same time, θ is the gain of sliding mode control switching item, which affects the dynamic characteristics and the system stability. When the external interference $f_d(x, t)$ increases, θ is also needed to increase to ensure the stability. However, the controller parameter θ is once determined, the online parameter adjustment cannot work then. In this paper, the parameter θ is optimized in real time based on reinforcement learning to solve the problem of adjusting parameters online.

Reinforcement learning is a kind of learning that maps from the environment to the action, which is to make the agent get the maximum cumulative return in the process of interacting with the environment. The reinforcement learning framework is shown in Fig. 2. The agent gets the initial state s_t from the environment perception and then take actor a_t. It will get the rewards back from the environment finally. In this way, a series of "state-action-reward" is generated. Until the end state, the Agent's purpose is to maximize the sum of reward values through continuous exploration of the environment and feedback.
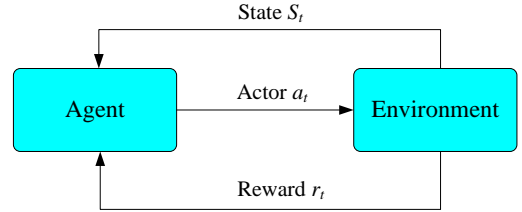
Fig. 2. The reinforcement learning framework

The improved deep intensive learning strategy combined with RBF neural network and Actor-critic (AC) learning method is utilized to optimize the parameter θ online. Meanwhile, the symbol function in the controller u_m will produce more chattering when θ increases. To remove the effect of chattering, the sat saturation function is selected instead of the sign function. Therefore, the controller u_m of the system becomes as follows.

$$u_m = -[g(\boldsymbol{\eta})]^{-1} \begin{bmatrix} f(\boldsymbol{\eta}) + c_1 \eta_2 + c_2 \eta_3 + c_0 e + (W + \hat{\theta}) sat(S) \\ + \kappa \left(c_1 e + c_2 \eta_2 + \eta_3 + c_0 \int_0^t e(t) dt \right) \end{bmatrix} \quad (13)$$

where, $\hat{\theta}$ is the actuating signal from learning RBF neural network based Actor-Critic.

$$sat(S) = \begin{cases} \dfrac{S}{\phi_0} & , \quad |S| \leq \phi_0 \\ \mathrm{sgn}(S) & , \quad |S| > \phi_0 \end{cases} \quad (14)$$

（13）

where, ϕ_0 is the boundary layer thickness.

In this paper, the universal approximation principle of RBF network is utilized to approximate the strategic function of Actor and the value function of Critic. Furthermore, Actor and Critic share the input layer and the hidden layer of RBF network, which reduces the storage space of the learning system. The structure block of control system is shown in Fig. 3.
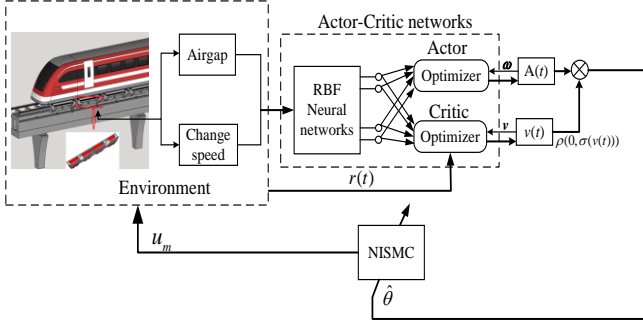
Fig. 3. The structure block of closed-loop control system

Input layer: Each input node of this layer is a component of the system state vector, which is as follows.

$$\boldsymbol{x}(t) = [\eta_1(t), \eta_2(t), \eta_3(t)]' \qquad (16)$$

Hidden layer: The output of the j neuron of the hidden layer can be expressed as follows.

$$\phi = \exp\left(-\frac{\|x(t) - c_j(t)\|^2}{b_j(t)}\right), \quad j = 1, \cdots, h \quad (17)$$

where, $c_j(t)$ is vector value of center point of the j node, $b_j(t)$ denotes the width parameter of the j node and h represents the number of neurons in the hidden layer.

Output layer: The following two expressions are used to calculate the strategic function of Actor and the value function of Critic.

$$A(t) = \sum_{j=1}^{h} \omega_j(t)\phi_j(t), \quad j = 1, \cdots, h \qquad (18)$$

$$v(t) = \sum_{j=1}^{h} v_j(t)\phi_j(t), \quad j = 1, \cdots, h \qquad (19)$$

where, $A(t)$ and $v(t)$ are the weights between the j node of the hidden layer and the node of the output layer respectively.

The output from Actor is not passed to the sliding mode controller directly, but superimposes a Gaussian search signal ρ, which is calculated as follows.

$$\hat{\theta} = A(t) + \rho(0, \sigma(v(t))) \qquad (20)$$

where, $\sigma(v(t)) = 1/(1 + e^{2v(t)})$.

Considering the influence of system error on control outputs, the reward function can be expressed as follows:

$$R(t) = \zeta r(t)$$

where, ζ is the enhancement signal coefficient of system error. $r(t)$ is the enhancement signal of system error, which can be defined as follows.

$$r(t) = \begin{cases} 1, & |e(t)| \le \varepsilon \\ 0, & \text{other} \end{cases} \qquad (21)$$

where, ε is the allowable error band. According to the sequential difference error δ_{TD} between adjacent state value functions, the learning performance index $E(t)$ of the system is defined as follows:

$$\delta_{TD}(t) = r(t+1) + \alpha v(t+1) - v(t) \qquad (22)$$

$$E(t) = \frac{1}{2}\delta_{TD}^2(t) \qquad (23)$$

where, α is the discount factor. The time series difference error δ_{TD} reflects the quality of the selected action

The network weight shown in the following forms is adjusted by the gradient descent method.

$$\omega_j(t+1) = \omega_j(t) + \xi_1 \frac{\partial E}{\partial \omega} \qquad (24)$$

$$v_j(t+1) = v_j(t) + \xi_2 \frac{\partial E}{\partial \omega} \qquad (25)$$

where, ξ_1 and ξ_2 are the learning rates for Actor and Critic respectively. $\frac{\partial E}{\partial \omega}$ is strategy gradient.

It is calculated by approximate strategy gradient estimation algorithm and combined with the equations (24) and (25).

$$\omega_j(t+1) = \omega_j(t) + \xi_1 \frac{\partial E}{\partial \delta_{TD}}\frac{\partial \delta_{TD}}{\partial \delta_\omega} \approx \omega_j(t) +$$

$$\xi_1 \frac{\partial (1/2\delta_{TD}^2)}{\partial \delta_{TD}}\frac{(\hat{\theta}-\theta)}{\sigma(v(t))}\phi(t) = \omega_j(t) + \xi_1\delta_{TD}\frac{(\hat{\theta}-\theta)}{\sigma(v(t))}\phi_j(t) \qquad (26)$$

The same can be found.

$$v_j(t+1) = v_j(t) + \xi_2\delta_{TD}\phi_J(t) \qquad (27)$$

The update algorithms of c_j and b_j in the RBF network are as follows.

$$c_j(t+1) = c_j(t) + \xi_3\delta_{TD}(t)\omega_J(t)\phi_j(t)\frac{x(t)-c_j(t)}{b_j^2(t)} \qquad (28)$$

$$b_j(t+1) = b_j(t) + \xi_4\delta_{TD}(t)\omega_J(t)\phi_j(t)\frac{\|x(t)-c_j(t)\|^2}{b_j^3(t)} \qquad (29)$$

where, ξ_3 and ξ_4 are the learning rates of the center point vector and the width respectively.

IV. NUMERICAL SIMULATION

In this section we constructed the dynamic model in MATLAB/Simulink environment to verify the effectiveness of the proposed approach. The parameters of the EMS maglev vehicle suspension system are listed in Table 1.

TABLE I.   PARAMETERS OF THE MAGLEV TRAIN SUSPENSION SYSTEM

Parameter name	Value	Parameter name	Value
Suspended mass m / kg	750	Air permeability $\mu_0 / (H \cdot m^{-1})$	$4\pi \times 10^{-7}$
Electromagnetic coil turns N_m	340	Magnetic leakage rate η	0
Electromagnetic pole area A_m / m^2	0.0196	Initial air gap	0.016
Inner resistance of solenoid coil R / Ω	1.2	Target air gap	0.008

The controller parameters are selected as follows:

$$c_1 = 100, \quad c_2 = 30, \quad c_0 = 37, \quad \kappa = 20, \quad W = 12,$$

$$c_j(0) = \begin{bmatrix} -2, -1, 0, 1, 2 \\ -2, -1, 0, 1, 2 \end{bmatrix}, \quad \zeta = 1.5, \quad \varepsilon = 0.001,$$

$$\xi_1 = \xi_2 = \xi_3 = \xi_4 = 0.4, \quad \alpha = 0.95, \quad \omega_j(0) = rands(5,1),$$

$$v_j(0) = rands(5,1), \quad b_j(0) = ones(1,5).$$

To further compare the superiority of the proposed control method based on reinforcement learning, we selected the traditional PID controller for comparison. The control parameters of PID are determined by trial and error as follows: kp=9500, kd=2100, ki=15. Note that our modeling is a voltage model, so the control object model is a third-order strong nonlinear differential equation, and no linearization is done.

The simulation result of PID controller is shown in Figure 4. The simulation results of the controller proposed in this paper are shown in Figure 5.

(a) Airgap response

(b) Change speed response
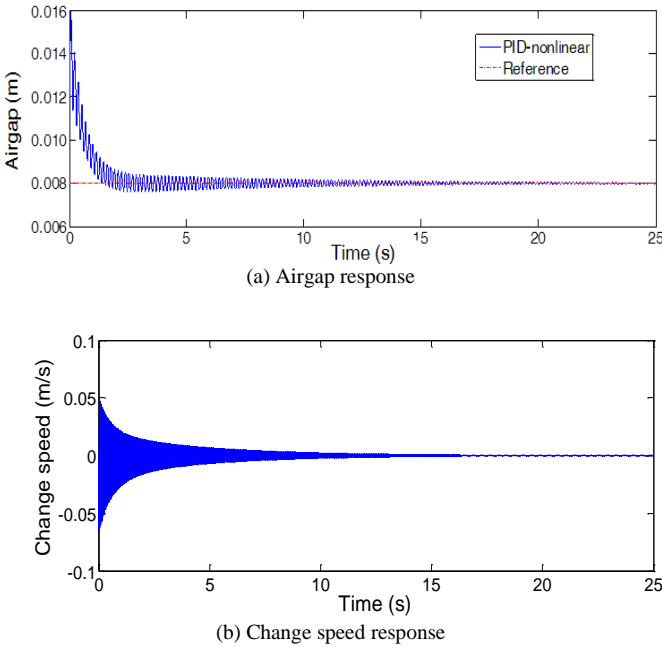
Fig. 4. Simulation results: Traditional PID controller under third-order nonlinear model

(a) Airgap response
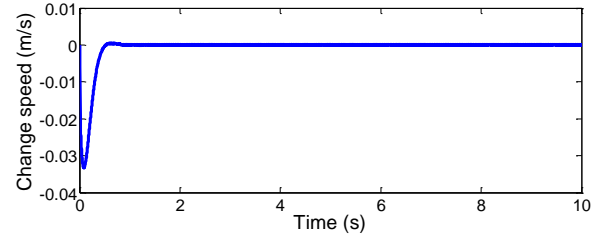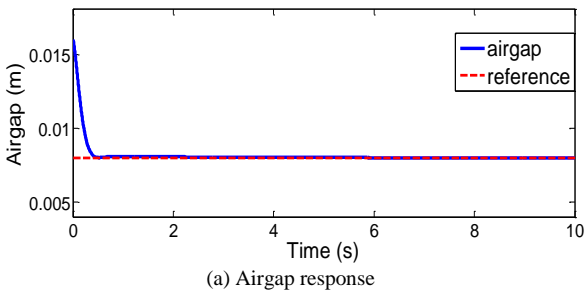
(b) Change speed response
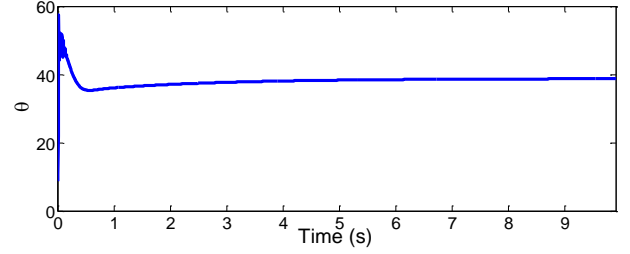
(c) The switching term gain θ with reinforcement learning

Fig. 5. Simulation results: the proposed control method based on reinforcement learning

In order to demonstrate the effects of reinforcement learning optimization, we also compared the proposed control method with traditional sliding mode controller [17]. The comparison results can be seen in the Table Ⅱ.

TABLE II. COMPARISON OF DIFFERENT METHODS

Method	Time	Static error	Max error	Chattering
PID	2 s	0.15 mm	0.52 mm	Bit
SCM	0.7 s	0.07 mm	0.09 mm	Yes
Proposed Controller	0.5 s	0.03 mm	0.04 mm	None

It can be seen from Fig. 4 that the controlled object is a third-order strong nonlinear equation, and the traditional linear PID control effect is poor, the rise time is about 2 s, and the overshoot is 6.3%. The traditional sliding mode controller has chattering phenomenon. The control approach based on AC learning method proposed in this paper has a significant improvement in control effect, with a rise time of about 0.5 s, an overshoot of 0.1%, and no chattering phenomenon.in the system state.

V. CONCLUSION

This article presents a nonlinear integral SMC method based on improved deep reinforcement Learning for the strong nonlinearity, open loop instability and susceptibility of maglev trains. Compared with traditional PID control, the system regulation time shorter, more adaptive and robust, and higher tracking accuracy. Simulation results demonstrate the effectiveness and superiority of this method.

This control strategy can also be applied to the intelligent control of nonlinear systems such as train active suspension system, track monitoring robot system and train speed control through appropriate improvement.

REFERENCES

[1] Luguang Y. Progress of high-speed Maglev in China. IEEE Transactions on Applied Superconductivity. 2002; 12(1): 944-947.

[2] Gottzein E, Meisinger R, Miller L. The" Magnetic Wheel" in the suspension of high-speed ground transportation vehicles. IEEE Transactions on Vehicular Technology. 1980; 29(1): 17-23.

[3] H. Wang, X. B. Zhang and G. Su. A new Maglev line system design and control strategy, Journal of Tongji University (Natural Science), 41 (7) (2013) 1112-1118.

[4] Chen C , Xu J , Lin G , et al. Fuzzy adaptive control particle swarm optimization based on T-S fuzzy model of maglev vehicle suspension system[J]. Journal of Mechanical Science & Technology, 2020, 34(1):43-54.

[5] Yong Wang, Hu Ding, Li-Qun Chen. Averaging analysis on a semi-active inerter–based suspension system with relative-acceleration–relative-velocity control. 2020, 26(13-14):1199-1215.

[6] Chen Chen, Xu Junqi, Rong Lijun, Pan Hongliang, Gao Dinggang. Nonlinear dynamics characteristics of maglev vehicle under track random irregularities[J].Journal of Traffic and Transportation Engineering,2019,19(04):115-124.

[7] Amare N D, Son Y I, Lim S.. Dual PIO-Based Controller Design for Robustness Improvement of a Magnetic Levitation System[J]. Journal of Electrical Engineering & Technology,2020,15(5): 1389–1398.

[8] Liu Wubin, Wang Quan. Research on model reduction of maglev control system under elastic track [J].Journal of National University of Defense Technology,2020,42(02):206-210.

[9] Li Yungang,Chen Hu,Zhang Xiao,Liu Hengkun. Research on Suspension and Guidance Technologies of EMS Maglev Trains Using V-shaped Guideway[J]. Journal of Tongji University (Natural Science),2012,40(11):1720-1724.

[10] Chen Xiaohao, Ma Weihua. Analysis on the Effect of Controller Time Delay on the Stability of Maglev System [J]. ELECTRIC DRIVE FOR LOCOMOTIVES,2019(02):139-143+147.

[11] Chen X S, Yang Y M. Adaptive PID Control Based on Actuator-evaluator Learning [J]. Control theory and application, 2011, 28 (8): 1187-1192

[12] Su L J, Zhu H J, Qi X H, Dong Hairui. Design of Four-rotor Height Controller Based on Reinforcement Learning[J]. Measurement and control technology, 2016, 35 (10): 51-53.

[13] Jing W, Paschalidis I C. An Actor-Critic Algorithm with Second-Order Actor and Critic[J]. IEEE Transactions on Automatic Control, 2017, PP(99):1-1.

[14] Su S X, Qi J P. Applied Research of Adaptive Online Planning Based on Reinforcement Learning [J]. Measurement and Control Technology, 2016,35(7): 124-127

[15] Sun Y, Xu J, Qiang H and Lin G. Adaptive neural-fuzzy robust position control scheme for maglev train systems with experimental verification [J]. IEEE Transactions on Industrial Electronics, 2019, 66(11): 8589-8599.

[16] Sun Y, Qiang H, Xu J, et al. Internet of Things-based online condition monitor and improved adaptive fuzzy control for a medium-low-speed maglev train system[J]. IEEE Transactions on Industrial Informatics, 2020, 16(4): 2629-2639.

[17] Sun Y , Li W , Lin G , et al. Dynamic Modeling and Nonlinear Control Research on Magnetic Suspension Systems of Low-speed Maglev Train[J]. Tongji Daxue Xuebao/Journal of Tongji University, 2017, 45(5):741-749.

Broadband Aperture Extension of A Passive Sonar Array Using Multi-Dimension Autoregression Model

Yuning Qian*
Nanjing Research Institute of Electronics Technology
Key Laboratory of IntelliSense Technology, CETC
Nanjing, China
inter101010@sina.com

Yawei Chen
Nanjing Research Institute of Electronics Technology
Key Laboratory of IntelliSense Technology, CETC
Nanjing, China
perse_1@163.com

Xinrong Cao
Nanjing Research Institute of Electronics Technology
Key Laboratory of IntelliSense Technology, CETC
Nanjing, China
caoxinrong@126.com

Jun Sun
Nanjing Research Institute of Electronics Technology
Key Laboratory of IntelliSense Technology, CETC
Nanjing, China
sunjun@ustc.edu

Abstract—This paper presents a broadband aperture extension approach of a passive sonar array based on multi-dimension autoregression (AR) model, integrated with failure element compensation. The multi-dimension AR model is proposed to build relations between broadband complex signals of array elements in frequency domain and predict the virtual element signals for wideband aperture extension, instead of the traditional time-domain AR model suitable for narrow-band aperture extension. In addition, the broken element compensation algorithm is designed to recover the failure array elements to ensure the feasibility and accuracy of AR modeling and virtual element prediction. Experimental studies verify the effectiveness of the presented approach for broadband array aperture extension.

Keywords—broadband aperture extension, multi-dimension AR model, failure element compensation

I. INTRODUCTION

With the development of shock absorption and noise reduction techniques, underwater target detection by passive sonar is becoming more and more difficult. In order to improve detection ability, passive sonar arrays are always required to have lager aperture, which can lead to higher array gain, signal to noise ratio(SNR) and azimuth resolution. However, due to the constraints of installation space and cost, the physical aperture of sonar array cannot be enlarged as expected. Therefore, the array extension approaches, which can make arrays obtain virtual aperture without increasing their physical size, have attracted noticeable attentions[1]. Covariance matrix extension(CME) and virtual array element estimation (VAEE) are two main kinds of array extension approaches. The CME approach uses matrix reconstruction [2,3] or matrix product [4,5] to obtain the extended covariance matrix for direction-of-arrival (DOA) estimation, but it cannot get extended array signals. On the other hand, the VAEE approach is able to estimate virtual array signals after extension so that it is suitable to serve as a pre-processing procedure for any array processing methods. For example, Friedlander[6] and Wang[7] proposes the virtual array transformation for array extension, which makes use of the relations between origin and extended steering vectors to transform the origin array into a larger virtual array. Nevertheless, this transformation does not take real signal features into consideration so that it can only be used for DOA estimation in a small bearing range. The linear prediction is another effective VAEE approach, which constructs time series models based on existing array signals to predict more virtual array elements. This approach has the advantages of simple structure, convenient construction and high accuracy, and it is widely used in array extension for the towed sonar [8-10], the horizontal fixed sonar [11], the planar sonar [12], the vector-array sonar [13] and the automotive radar [14]. However, these proposed linear prediction approaches usually focus on narrow-band array extension, while acoustic signals emitted by underwater targets are mainly broadband signals. Hence, broadband array aperture extension based on linear prediction approaches should be further investigated.

Inspired by the prior studies and discussions, this paper presents a broadband aperture extension approach for a passive sonar array. Instead of using traditional time-domain AR models for narrow-band array extension, the multi-dimension AR model is firstly proposed to build relations between broadband complex signals of array elements in frequency domain, and predict the virtual element signals for wideband aperture extension. Then the broken element compensation algorithm is designed to recover the failure array elements to ensure the feasibility and accuracy of AR modeling and virtual element prediction. Finally, the extended array signals can be further utilized for beamforming and DOA estimation. The organization of the rest paper is as follows. Section II introduces the principle of broadband array extension using multi-dimension AR model. Section III gives the results of experimental studies, together with some discussions. The conclusions are finally stated in Section IV.

II. METHODOLOGY

A. Problems

For a N-element uniform line array with sensor spacing d and a far field target signals from azimuth θ, the received target signals of each array element can be written as:

$$X(t) = [x_1(t), ..., x_n(t), ..., x_N(t)]$$
$$= [x_1(t), ..., x_1(t - \tau_n), ..., x_1(t - \tau_N)]$$
(1)

It can be noted that the received signals of nth array element is the time delay result of 1st element and the delay value $\tau_n = (n-1)d\sin\theta/c$, where c represents the sound speed. To extend the array aperture by linear prediction approach, an AR model can be constructed using $[x_1(t)... x_N(t)]$ to predict $x_{N+1}(t)$ with the potential condition that the series $[x_1(t)... x_N(t)]$ is correlated. The condition is satisfied when the target signal is narrow-band, while the condition is possibly unsatisfied when the target signal is broadband. Generally speaking, the signals of underwater acoustic target are always broadband noise, e.g. Gaussian noise. Mathematically, if $x_1(t)$ is Gaussian noise, the correlation coefficient between $x_1(t)$ and $x_1(t-\tau_n)$ is 0 when $\tau_n \neq 0$. Under this circumstance, AR model construction is most likely ineffective for broadband array signals $X(t)$ in time domain. Therefore, $X(t)$ is considered to be transformed into frequency domain and equation (1) can be rewritten as:

$$X(f) = [x_1(f), ..., x_n(f), ..., x_N(f)]$$
$$= [x_1(f), ..., x_1(f)e^{-j2\pi f\tau_n}, ..., x_1(f)e^{-j2\pi f\tau_N}]$$
$$= [x_1(f), ..., x_1(f)e^{-j2\pi f\frac{(n-1)d\sin\theta}{c}}, ..., x_1(f)e^{-j2\pi f\frac{(N-1)d\sin\theta}{c}}]$$
(2)

where $x_n(f)$ is the Fourier transform result of $x_n(t)$ and c is the sound speed. Through Fourier transform, the signal amplitude of each array element in a certain f is the same, and there is only a fixed phase difference $e^{-j2\pi fd\sin\theta/c}$ between any adjacent array element. Hence, for each frequency point, $[x_1(f)... x_N(f)]$ is correlated so that AR model can be established to predict more virtual array element signals.

However, the proposed frequency broadband array extension approach still has two main problems. Firstly, the frequency array signals $X(f)$ are complex signals and traditional AR model is a time-series model suitable for real signals. That means traditional AR model cannot be directly used for frequency signal modeling and prediction. Secondly, in real case, sonar arrays are often worked in hostile marine environment and array element failures are inevitable. Once some array elements are broken, the phase difference between the adjacent health elements is no longer a fixed value and AR model cannot be constructed using these non-uniform array signals to predict a series of uniform virtual elements.

B. Complex AR Model Construction

The AR model constructed by frequency array signals $X(f)$ can be defined as:

$$x_{N+1}(f) = \sum_{j=1}^{p} A_j x_{N+1-j}(f) + \varepsilon_{n+1}$$
(3)

where A_j is the model parameter, p is the model order and ε_{n+1} stands for the white noise. It can be seen that the virtual element signals $x_{N+1}(f)$ can be predicted by several existing element signals $x_{N+1-p}(f)...x_N(f)$ using AR model in each frequency point $f \in [f_1, f_d]$. However, traditional real AR model cannot be directly constructed by complex data $x_n(f)$. To solve this problem, the complex signal $x_n(f)$ can be converted to a vector x_n consisting of the signal real part $x_{n,r}(f)$ and signal imaginary part $x_{n,i}(f)$, and then equation (3) can be rewritten as a two-dimension AR model:

$$x_{N+1} = \sum_{j=1}^{p} A_j x_{N+1-j} + \varepsilon_{n+1}$$
(4)

$$\begin{bmatrix} x_{N+1,r}(f) \\ x_{N+1,i}(f) \end{bmatrix} = \sum_{j=1}^{p} \begin{bmatrix} A_j(1) \\ A_j(2) \end{bmatrix} \begin{bmatrix} x_{N+1-j,r}(f) \\ x_{N+1-j,i}(f) \end{bmatrix} + \begin{bmatrix} \varepsilon_{n+1,r} \\ \varepsilon_{n+1,i} \end{bmatrix}$$
(5)

where $x_n \in R^{2*1}$ and $A_j(i) \in R^{1*2}$. To estimate the model parameter A_j, the least square parameter estimation method for multidimensional AR model proposed in reference [12] is utilized in this paper.

The observed vector $L \in R^{(N-p)*2}$, sample matrix $B \in R^{(N-p)*2p}$ and parameter matrix $\Phi \in R^{2p*2}$ can be defined as:

$$L = \begin{bmatrix} x_N^T \\ x_{N-1}^T \\ \vdots \\ x_{p+1}^T \end{bmatrix}, B = \begin{bmatrix} x_{N-1}^T & x_{N-2}^T & ... & x_{N-p}^T \\ x_{N-2}^T & x_{N-3}^T & ... & x_{N-p-1}^T \\ \vdots & \vdots & \vdots & \vdots \\ x_p^T & x_{p-1}^T & ... & x_1^T \end{bmatrix}, \Phi = \begin{bmatrix} A_1^T(1) & A_1^T(2) \\ A_2^T(1) & A_2^T(2) \\ \vdots & \vdots \\ A_p^T(1) & A_p^T(2) \end{bmatrix}$$
(6)

where T is the transposition symbol. Based on the mathematical inference in [15], the least square estimation of model parameters can be calculated as:

$$\Phi = (B^T B)^{-1} B^T L$$
(7)

Particularly, in order to using equation (7), model order p need to be chosen firstly. According to the matrix theory:

$$2p = r(B^T B) \leqslant r(B) \leqslant \min(2p, N-p)$$
(8)

where $r(\bullet)$ is the rank of the objective matrix. To ensure the full rank and invertibility of $B^T B$, $N-p$ should be larger than $2p$ so that p should be smaller than $N/3$. In this study, p is chosen as $N/3$ to confirm the effectiveness of equation (7). Using Φ and equation (8), the vector x_{n+1} can be obtained and virtual element signals $x_{N+1}(f)$ can be further calculated as:

$$x_{N+1}(f) = x_{N+1,r}(f) + i \cdot x_{N+1,i}(f)$$
(9)

Then, more virtual element signals $x_{N+2}(f), ..., x_{N+k}(f)$ can be predicted as the same procedure.

C. Broken Array Element Compensation

In real working environment, element failure for sonar arrays is inevitable. From equation (3) to (6), it should be noted that AR model need to be constructed based on signals of uniform array elements. Hence, if element failure occurs, the broken element signals should be recovered to build a completely uniform array. Inspired by array extension approach using complex AR model in section A and B, a real-time broken element compensation algorithm based on complex partial AR model (pAR) for a N-element uniform line sonar array is proposed in this study. The algorithm procedure is introduced as follows:

1) The location numbers of broken elements are obtained.

2) Each separated broken element or each continuous broken element part is identified as a broken segment shown in Fig. 1, $i=1,2,...,k$. The broken array number in each segment is labeled as m_i.

3) From left to right of the array, the element numbers of healthy segments on the left and right side of each m_i are computed as h_i and h_{i+1}. If $\max(h_i, h_{i+1}) \geq 3m_i$, the element signals in the left or right healthy segment corresponding to the maximum value of h_i and h_{i+1} are utilized to establish an

AR model as illustrated in section 2.2 and predict the element signals of segment m_i. On the other hand, if max(h_i, h_{i+1})<2m_i, the elements after compensation in broken segment m_{i-1} are treated as healthy elements, and then all element signals involved in h_{i-1}, h_i and m_{i-1} are utilized to estimate element signals for segment m_i.
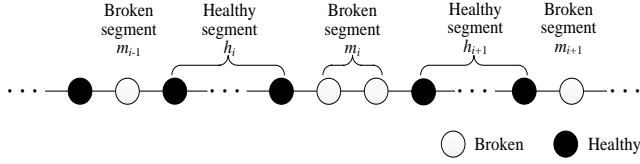
Fig. 1. Schematic drawing illustrating the basic idea of broken element compensation

D. Framework of Broadband Array Extension

By combining complex AR model with broken element compensation algorithm, a broadband aperture extension approach is proposed and the procedure shown in Fig. 2 can be summarized as follows:

1) Firstly, the array signals in the interested frequency band are extracted and transformed into frequency domain.

2) Secondly, broken element compensation algorithm is utilized to recover the signals of each failure element.

3) Thirdly, for different frequency points, different complex AR models are constructed using existing array element signals to estimate more virtual element signals. As a result, signals of an extended array with larger aperture are obtained.

4) Fourthly, the extended array signals are used for beamforming and DOA estimation.
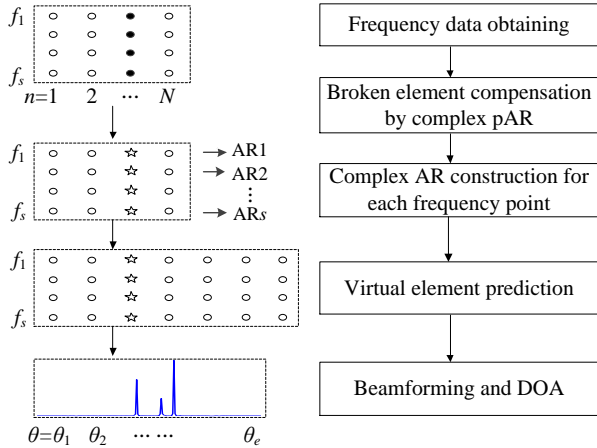
Fig. 2. Flowchart of broadband array extension approach

III. EXPERIMENTAL STUDY

In order to verify the effectiveness of the proposed broadband aperture extension approach, the simulation test and actual experiment are both conducted.

A. Simulation Test

There are three simulation targets (target 1 to target 3) in the space and the target signals are broadband Gaussian noise ranging as 500-1000Hz. The target bearings are -10, 10 and 16 degree, and the SNRs are -6dB, -6dB and -20dB respectively. The received array is a 64-element uniform line array with sensor spacing 0.75m.

To begin with, the initial array is extended to a 128-element array by the traditional time-domain AR extension method and the proposed frequency-domain complex AR extension method, and DOA results of initial array and extended arrays using conventional beamforming (CBF) are shown in Fig. 3. It can be seen from Fig. 3 that the DOA amplitudes of target 1 and target 2 computed by frequency-domain extension array are 3dB higher than that computed by initial array, and the DOA amplitude of weak target 3 computed by frequency-domain extension array are 1.5dB higher than that computed by initial array. On the other hand, the DOA amplitudes calculated by time-domain extended array is almost the same as that calculated by initial array and there even exists a false target in its DOA result. This means time-domain extension is nearly ineffective for wideband array signals in this study, which can be explained as illustrated in section 2.1. The comparison results indicate that the proposed frequency-domain extension method is more beneficial than the traditional time-domain extension method.

Then, 25% array elements, marked as 14, 18:19, 22, 31, 37:38, 51:59, are assumed to be broken. The DOA results of the initial broken array, the extended array with broken compensation and the extended array without broken compensation are calculated by CBF and shown in Fig. 4. It can be noted that the DOA amplitudes of extended array with broken recovery is higher than that of extended array without broken recovery, especially for target 1 and 2. Furthermore, there also exists several false targets for extended array without compensation, which is the sidelobe expansion resulting from the broken elements. Therefore, the proposed complex pAR compensation method is effective for broken array recovery and DOA estimation.

In addition, to further verify the applicability of the proposed aperture extension approach, the wideband adaptive beamforming(ABF) [16] and MUSIC algorithm are applied to the extended array signals for DOA estimation, and the results are shown in Fig. 5 and 6. From these two figures, it can be seen that ABF and MUSIC algorithms applied to extended arrays can significantly improve the target SNR and weak target resolution, compared with those applied to initial arrays.
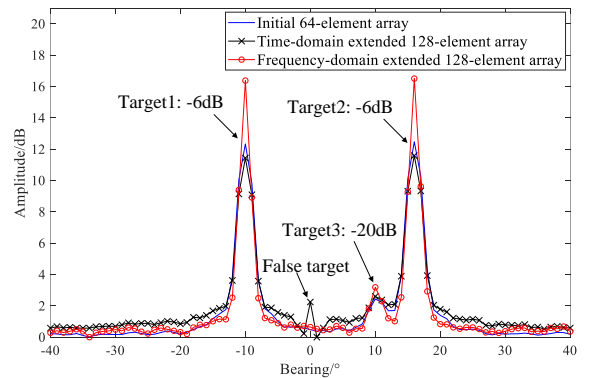
Fig. 3. DOA results of time-domain and frequency-domain extended array using CBF
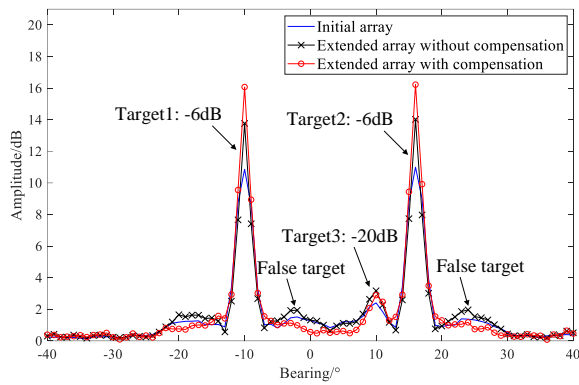
Fig. 4. DOA results of extended arrays with and without compensation using CBF
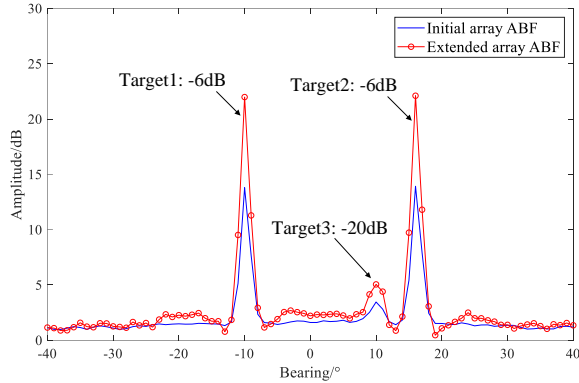
Fig. 5. DOA results of extended arrays using ABF
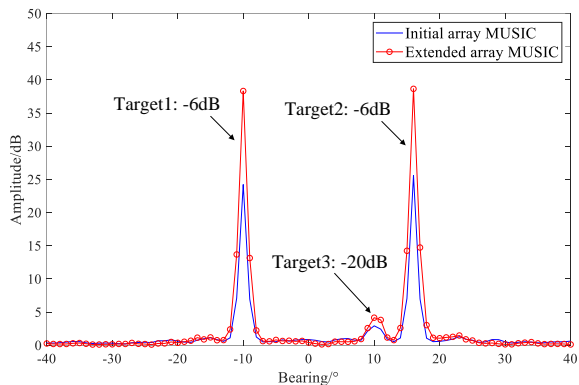
Fig. 6. DOA results of extended arrays using MUSIC

Fig. 7. BTR results of sonar signals

Fig. 8. DOA results in time 28s

B. Actual Experiment

At last, real array signals sampled from a 64-element sonar in a sea experiment are analyzed to confirm the above analysis results. The source is located in bearing -25° and emits 500-1000Hz wideband Gaussian noise. Using the proposed broadband aperture extension approach, the initial sonar array is extended to a 128-element array, and the bearing-time-recording (BTR) and DOA results in time 28s are computed by CBF and shown in Fig. 7 and 8. It can be noted from Fig. 7 that the beam width of the source obtained by the extended array are narrower that obtained by the initial array. Furthermore, it can be seen from Fig.8 that the SNRs of all targets gained by the extended array are significantly improved compared with those gained by the initial array. Therefore, the extended array has the better bearing resolution and SNR improvement abilities.

IV. CONCLUSIONS

A broadband aperture extension approach based on multi-dimension AR model for a passive sonar array is presented in this paper. The frequency-domain multi-dimension AR model is proposed for broadband array signal modeling and virtual element generation, combined with the broken element compensation algorithm for failure element recovery. The experimental studies on simulation data and real sonar data indicate that the proposed broadband aperture extension approach can significantly improve the target SNR and the bearing resolution ability. The proposed complex pAR compensation method is effective for broken array recovery, and the frequency-domain complex AR model is more beneficial than traditional time-domain AR model for wideband array aperture extension. Furthermore, the proposed broadband aperture extension approach is suitable to serve as a pre-processing procedure for various DOA methods, such as CBF, ABF and MUSIC.

REFERENCES

[1] B. Kim, S. Kim, and J. Lee, "A novel DFT-based DOA estimation by a virtual array extension using simple multiplications for FMCW radar," Sensors, vol. 18, no. 5, pp. 1560, 2018.

[2] F. Xing, C. Zhou, Y. Gu, et al., "Toeplitz matrix reconstruction of interpolated coprime virtual array for DOA estimation," Proceedings of Vehicular Technology Conference, Sydney, Australia, 2017.

[3] F. Yan, X. Yan, J. Shi, et al., "MUSIC-like direction of arrival estimation based on virtual array transformation", Signal Processing, vol. 139, pp. 156-164, 2017.

[4] Y. Dong, C. Dong, W. Liu, et al., "2-D DOA estimation for L-shaped array with array aperture and snapshots extension techniques," IEEE Signal Processing Letters, vol. 24, no. 4, pp. 495-499, 2017.

[5] H. Yamada, S. Shirai, and Y. Yamaguchi, "DOA estimation of wideband signals using extended virtual array," Proceedings of 2014 3rd Asia-Pacific Conference on Antennas and Propagation, Harbin, China, pp. 802–805, 2014.

[6] B. Friedlander, "Direction finding using an interpolated array," Signal Processing Technology, vol. 3, no. 5, pp. 2951-2954, 1990.

[7] Y. Wang，H. Chen, and S. Wan．"An effective DOA method via virtual array transformation," Science in China(Series E: Technological Sciences) , vol. 29, no. 6, pp. 518-524, 1999.

[8] S. Swingler and R. Walker, "Line-array beamforming using linear prediction for aperture interpolation and extrapolation," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 1, pp. 16-30, 1989.

[9] S. Stergiopoulos and N.T. Allcott, "Aperture extension for a towed array using an acoustic synthetic aperture or a linear prediction method," Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, San Francisco, USA, 1992.

[10] Y. Hou, J. Huang and W. Shi, "Linear prediction based ETAM method for underwater target bearing estimation", System Engineering and Electronics, vol. 32, no. 9, pp. 1803-1810, 2010.

[11] P. Hu，Y. Yang, and S. Yang, "Virtual array beamforming based on linear prediction," Technical Acoustics, vol. 26 no. 4, pp. 714 -717, 2007.

[12] H. Xiong, B.Yuan, W. Hu, et al., "Precise bearing method of building virtual array with the improved linear prediction algorithm," Journal of National University of Defense Technology, vol. 37, no. 1, pp. 142-147, 2015.

[13] M. Wang and W. Zhu, "Virtual arrays' beamforming of non-equidistan vectors array", Audio Engineering, vol. 36, no. 4, pp.42-45, 2012.

[14] H. Sim, S. Lee, S.Kang, et al., "Enhanced DOA estimation using linearly predicted array expansion for automotive radar systems," IEEE Access, vol. 7, pp. 47714-47727, 2019.

[15] Y. Qian, R. Yan, and R.X. Gao, "A multi-time scale modeling approach to remaining useful life prediction in rolling bearings," Mechanical System and Signal Processing, vol. 83, pp. 549-567, 2017.

[16] S.D. Somasundaram, "Wideband robust capon beamforming for passive sonar," IEEE Journal of Oceanic Engineering, vol. 38, no. 2, pp. 308-322, 2013.

Improved Rotating Speed Estimation and Bearing Fault Diagnosis Using Multi-Channel Vibration Signals

Xiaoxian Wang
Traction Power State Key Laboratory
Southwest Jiaotong University
Chengdu, China
wxx612@mail.ustc.edu.cn

Jiahao Niu
College of Electrical Engineering and
Automation
Anhui University
Hefei, China
niu.jiahao@qq.com

Siliang Lu*
Traction Power State Key Laboratory
Southwest Jiaotong University
Chengdu, China
lusliang@mail.ustc.edu.cn

Fang Liu
College of Electrical Engineering
and Automation
Anhui University
Hefei, China
ufun@foxmail.com

Yongbin Liu
College of Electrical Engineering
and Automation
Anhui University
Hefei, China
lyb@ustc.edu.cn

Abstract—A rotation speed estimation algorithm for motor bearing fault diagnosis based on multi-channel vibration signals analysis under variable speed conditions is proposed in this study. First, the vibration signals are acquired by a tri-axial accelerometer installed on the motor housing. Multiple instantaneous frequencies (IFs) are extracted from the vibration signals using time-frequency analysis, and then multiple rotating angle curves are calculated based on the IFs. The rotating angle curves are selected and fused into a new one with high accuracy. The original vibration signal is resampled in accordance with the accumulated rotating angle curve. Bearing fault type can be determined from the spectrum of the demodulated resampled vibration signal. The method proposed can realize rotating speed estimation and fault diagnosis using only the vibration signals. Hence, it provides a new solution to realize bearing fault diagnosis under speed variation conditions without using a tachometer.

Keywords—motor bearing fault diagnosis; time-frequency analysis; signal resampling; multi-channel signal fusion

I. INTRODUCTION

Bearing is one of the most pivotal components in a motor or a generator. Generally, a pair of bearings are installed on two ends of a shaft to support its spinning. The bearings are subjected to both the radial and axial stresses, and hence the bearing are prone to failure after long time of rotations. The healthy condition of the bearing can be monitored by analyzing the signals from the sensors installed on the motor housing. Accelerometer, acoustic sensor, current sensor, temperature sensor and magnetic sensor are commonly used for bearing condition monitoring and fault diagnosis. Among them, vibration signal acquired from an accelerometer is sensitive to the variation of the bearing conditions.

Hence, vibration signal processing methods have been extensively investigated for bearing fault diagnosis [1, 2]. Generally, an accelerator is installed on a proper place of a motor to acquire the vibration signal. When a localized fault occurs in the motor bearing, the vibration signal will transmit to the sensor via a complex path. The quality of the acquired vibration signal is highly related to the sensor installation and noise interference. In only one channel of vibration signal is processed, the extracted features may not be complete and then the accuracy of fault diagnosis is affected.

The information fusion technique that fuses multiple sensor signals can improve the performances of feature extraction and fault diagnosis [3, 4]. By fusing different signals, the complementary information is enhanced, and the noise level is averaged and attenuated at the same time. Vibration signals fusion methods have been applied to bearing fault diagnosis. For instance, a multi-sensor feature fusion method combining sparse autoencoder and deep belief network was proposed for bearing fault diagnosis [5]. A tri-axial vibration information fusion algorithm was investigated to realize fault diagnosis in the conditions of speed variation [6].

When a rotating machine works in constant speed condition, the fault-induced features will appear periodically in the sampled signal. The time-domain or frequency-domain signal analysis technics can be utilized to estimate the intervals or frequencies of the periodic features, and then machine fault can be diagnosed by comparing the estimated features with the theoretical ones. However, when the machine speed varies with time, the signal features become non-stationary and the traditional frequency-domain analysis methods become inapplicable.

Order tracking technique can be used to process the time-varying signals by resampling the original signals on angular-domain. The rotation speed information, which is always obtained from a tachometer, is required to conduct order analysis. However, the speed information is not available in some cases, such as the driving motor is controlled using sensor-less strategy. On this occasion, the rotating speed can be estimated from other signals such as video stream, motor current and vibration signal [7].

When the rotation speed is estimated from video stream, the speed can be obtained through a contactless and in noninvasive way. However, an expensive high-speed camera is required, and processing the video stream requires a high volume of computations and storage spaces. The estimation accuracy is also easily affected by circumstance factors such as illumination, background color and distance between the camera and rotation part. Generally, the frequency of the motor phase current is highly relative to that of the motor rotor, especially for the synchronous motors. However, an extra current sensor is necessary to obtain the current signal. In some scenarios, acquisition of the currents of the high-

---

*Siliang Lu is the corresponding author.
(e-mail: lusliang@mail.ustc.edu.cn).

voltage motors is hazardous potentially. As the vibration signal has been acquired for bearing fault diagnosis, if the rotation speed can be directly estimated from the vibration signal, the measurement and instrument system will be more simple and efficient. However, the vibration signal is always corruptly by background noise. Hence, improving the angle estimation accuracy from the vibration signal is still a challenge.

To address this issue, a method that combines multi-sensor signal analysis and order tracking is investigated for motor bearing fault diagnosis under conditions of speed variation in this study. The vibration signals from a tri-axial accelerometer are acquired and processed using a time-frequency analysis method. The instantaneous frequencies (IFs) are extracted and the accumulated rotating angle curves of the motor are computed from the IFs. The rotating angle curves are selected and fused into a new one. In accordance with the fused rotating angle, the vibration signal is resampled on angular domain. The spectrum of the demodulated resampled signal is computed to reveal the fault indicators. The proposed method shows potential applications in diagnosis of motor bearing faults under variable speed conditions using only vibration signals.

The rest of the paper is organized as follows. Sec. II introduces the proposed rotating speed estimation and bearing fault diagnosis method. Sec. III introduces the experimental setup for signal acquisition. Sec. IV evaluates the performance of speed estimation using tri-axial accelerometer and verifies the proposed method's effectiveness for bearing fault detection. The conclusions are provided in Sec. V.

II. ROATATING SPEED ESTIMATION AND FAULT DIAGNOSIS

In this study, multiple IFs are extracted from the tri-axial vibration signals, and a high-accurate rotating angle curve is fused from the multiple signals for order tracking and fault diagnosis. The detailed steps are introduced in the follows.

A. Multiple IFs Extraction using Wavelet Synchrosqueezed Transform (WSST)

The vibration signal under variable speed condition is time-varying and non-stationary. Time-frequency analysis technology is used to extracted the IFs from the time-varying vibration signals. Many time-frequency analysis algorithms have been investigated for charactering the vibration signals. In this study, the wavelet synchrosqueezed transform (WSST) is used for time-frequency analysis as this method has high accuracy for IF estimation and extraction. A continues vibration signal with J components is expressed as [8]:

$$V(t) = \sum_{i=1}^{J} A_i(t)\cos\omega_i(t) + n(t) \quad (1)$$

in which $\omega_i(t)$ and $A_i(t)$ are the IF and instantaneous amplitude (IA) of the ith component of the vibration signal, respectively, and $n(t)$ is the measured noise. $V(t)$ in (1) is a time-varying signal, the WSST extracts the components' IFs and IAs and then reconstructs the components. Next, the procedures of IF extraction are introduced.

The IFs are estimated using the continues wavelet transform as:

$$W_V(a,b) = a^{-1/2} \int V(t)\psi^*(\frac{t-b}{a})dt \quad (2)$$

where b and a are the shift and scale coefficients of the wavelets, $\psi(\cdot)$ represents the mother wavelet and $\psi^*(\cdot)$ represents its conjugation. The IF of the vibration signal can be computed as:

$$f_V(a,b) = \frac{-i\partial(W_V(a,b))}{W_V(a,b)\partial b} \quad (3)$$

Eq. (3) builds a mapping relationship from (b, a) to $(b, f_V(a, b))$. The WSST redistributes the energy on the time-scale plane, and the energy on time-frequency plane is converted as:

$$T_V(f,b) = \int_{\{a:W_V(a,b)\neq 0\}} W_V(a,b)a^{-3/2}\frac{1}{\delta}h\left(\frac{f-f_V(a,b)}{\delta}\right)da \quad (4)$$

in which δ represents the accuracy and $h(\bullet) \in C_0^\infty$ represents a smooth function meeting $\|h\|_{L^1} = 1$. Subsequently, ridge extraction algorithm is used to extract the IFs and the corresponding IAs from the time-frequency distributions of the vibration signals.

B. Rotating Angle Curve Calculation

The signal components corresponding to the IFs and IAs are reconstructed as:

$$G_r[n] = x_m[n] \times cosm[n] + y_m[n] \times sinm[n], n = 0,1,...,N-1, \quad (5)$$

where $G_r[n]$, $r = 1,2,...,J$, is the rth signal component, N is the number of sampling points of the discrete signal, $x_m[n]$ and $y_m[n]$ are the two quadrature components of the IAs, and $cosm[n]$ and $sinm[n]$ can be expressed as:

$$\begin{cases} cosm[n] = \cos\left(2\pi\left(cumtrapz\left(t[n], IF[n]\right)\right)\right) \\ sinm[n] = \sin\left(2\pi\left(cumtrapz\left(t[n], IF[n]\right)\right)\right) \end{cases} n = 0,1,...,N-1 \quad (6)$$

where $cumtrapz()$ is the numerical integration function using cumulative trapezoidal method.

The accumulated rotating angle is then calculated using the Hilbert transform and subsequently normalized with respective to the base rotation frequency. To improve the computational efficiency, 3 IF curves with the highest energies are extracted from each channel of vibration signal. Finally, total 9 normalized rotating angle curves from x-, y- and z-axis can be obtained as:

$$\Phi = \{RA_i^j\}, i = x, y, z, j = 1, 2, 3 \quad (7)$$

C. Selection and Fusion of the Rotating Angle Curves for Fault Diagnosis

Considering that the mechanical system is an inertial system and thus the IFs should have similar variation trends. Given this, the 9 rotating angle curves are selected and fused into a new one to reduce the effects of noise interference. In this study, correlation analysis is used to select the curves. If two IF curves are similar, their cross correlation coefficient is larger, and vice versa. The maximal and minimal correlation coefficient values are 1 and 0, respectively. The

correlation coefficients among the set in (7) are computed as shown below:

$$CC = \begin{bmatrix} Corr\left(RA_x^1, RA_x^1\right) & ... & ... & Corr\left(RA_x^1, RA_z^3\right) \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ Corr\left(RA_z^3, RA_x^1\right) & ... & ... & Corr\left(RA_z^3, RA_z^3\right) \end{bmatrix}$$

(8)

in which $Corr()$ is the function for calculating the correlation coefficients between two elements. The correlation coefficients among one IF curve and other 8 IF curves are calculated and listed in each column or each row. The averaged value for each column/row is calculated. Next, 5 rotating angle curves with the highest averaged correlation coefficient values are selected for fusion. The discrete fused rotating angle curve is denoted as $RA_f[n]$, $n = 1,2,…, N$. Then, the original bearing vibration signal is resampled for fault diagnosis as:

$$RV[n] = \text{resample}(V[n], RA_f[n]) \quad (9)$$

in which $V[n]$ is the discrete form of $V(t)$ and resample() is an embedded function in MATLAB. The original vibration signal is sampled at fixed sampling frequency, namely, the sampling intervals are the same. By using the signal resampling algorithm, the time-domain signal is resampled according to a vector with equal angle increment. Afterward, the amplitude spectrum of $RV[n]$ is computed via fast Fourier transform. If a localized fault exists in the tested bearing, it is expected to find the distinct frequency components relative to the fault type in the spectrum.

III. EXPERIMENTAL SETUP

To acquire the vibration signals for analysis, an experimental setup is designed. An outer raceway fault is set on the tested bearing using electric spark machining. The fault bearing is installed on a permanent magnet synchronous generator (PMSG).
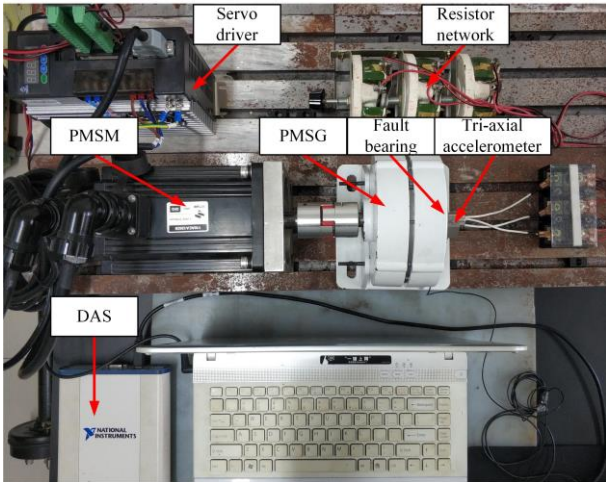
Fig. 1. Experimental test rig.

In practice, the PMSG is usually used in a direct-drive wind turbine. Because the wind speed and direction are always varied with time and meteorological condition, the spinning speed of the PMSG shaft changes with time. Hence, bearing fault detection in PMSG should consider the conditions of speed variation. A tri-axial accelerometer with type of 1A302E (DONGHUA Inc.) is mounted on the surface of the PMSG through a screw. The sensitivity of the accelerometer is 10 mV/g. The PMSG is driven by a permanent magnet synchronous motor (PMSM) through a mechanical coupling, and the motor speed is set variable with time through a servo driver.

A resistor network connects to the three-phase output of the PMSG as the electrical load. The three-channel vibration signals are acquired by a data acquisition system (DAS) with type of NI USB4432. The encoder signal from the PMSM is also acquired as the reference signal to assess the rotating angle estimation accuracy. The sampling frequency of the DAS is configured as 25.6 kHz and the duration of signal sampling is configured as 4 s. The tri-axial sensor signals are transmitted to a computer through a high-speed USB interface for analysis. The bearing parameters are shown in Table I. The fault characteristic order of the bearing with outer raceway fault (denoted as FCO$_O$) is computed as 3.055 in accordance with the bearing parameters.

TABLE I. PARAMETERS OF THE PMSG BEARING

Bearing type	Outside diameter	Inside diameter	Number rollers	FCO$_O$
6203	40 mm	17 mm	8	3.055

IV. VERIFICATION OF THE MULTI VIBRATION SIGNAL FUSION METHOD

The multi-channel vibration signals acquired from the experimental setup are processed to validate the proposed method's performance. The vibration signals' waveforms from different axes are shown in Fig. 2(a). The amplitudes and variation trends of the signals from different axes are similar. The amplitude spectra and the envelope spectra are shown in Figs. 2(b) and 2(c), respectively. As the bearing spins at variable speed, the amplitude and envelope spectra cannot reveal useful information for diagnosis of bearing fault. From another aspect, it can be seen from the spectra that the frequency components in x-axis contain more information than those in y-axis and z-axis.

The rotating angle curves are calculated from the extracted 9 IFs using the method introduced in Sec. II.B. The rotating angle curves are normalized with respect to the base rotating speed and the results are shown in Fig. 4. The angle curves have obvious differences due to the errors induced from the IF extraction procedure. Subsequently, 5 rotating angle curves with the highest accuracies are selected from the 9 curves using the correlation analysis method introduced in Sec. II.C. The correlation coefficients among the curves are calculated and the corresponding curves with the top 5 average coefficients are selected as the final curves. In the calculated coefficient matrix CC in Eq. (8), the average values are calculated for each row and the results are shown in Fig. 5. It can be found that the curves 1, 2, 5, 8 and 9 have the highest scores, and the curves 6 and 7 have the lowest scores.

The 5 curves are then fused into a new curve $RA_f[n]$. The fused accumulated rotation angle is used to resample the original vibration signal at x-axis and the resampled signal is displayed in Fig. 6(a). Some impulses induced by bearing faults with quasi-equal intervals can be seen in the waveform after angular-domain resampling. The spectrum of the

509

demodulated resampled signal is shown in Fig. 6(b). The frequency components FCO_O and its $2\times$ harmonic are clearly displayed in the spectrum. This result indicates that an outer raceway fault occurs in the PMSG bearing, thereby validating the proposed method's effectiveness.
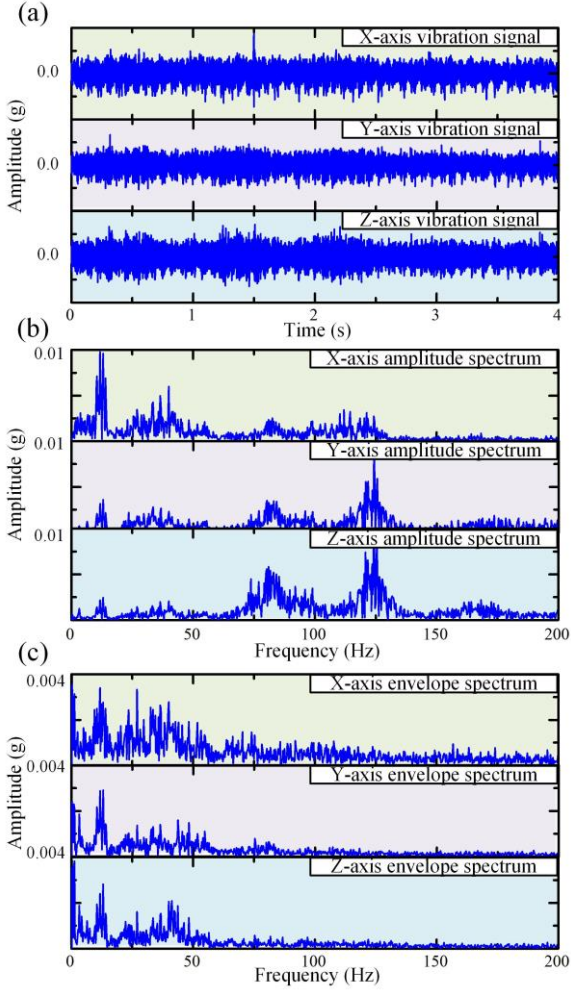
Fig. 2. (a) Tri-axial vibration signal, (b) vibration signal spectrum, and (c) envelope spectrum.
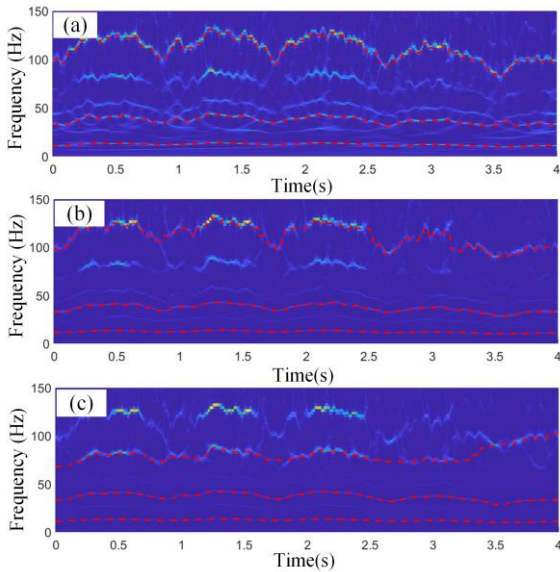
Fig. 3. Time-frequency images of the (a) x-axis (b) y-axis, and (c) z-axis vibration signal.
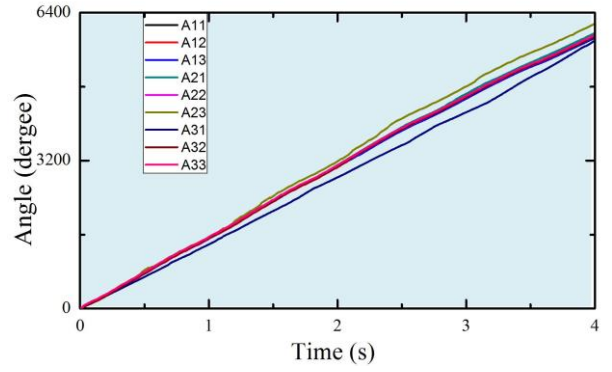
Fig. 4. The extracted nine normalized rotating angle curves.
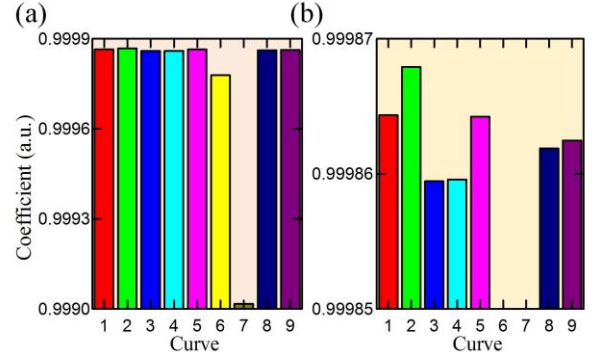
Fig. 5. (a) Average values of correlation coefficients of the curves, and (b) zoom-in figures.
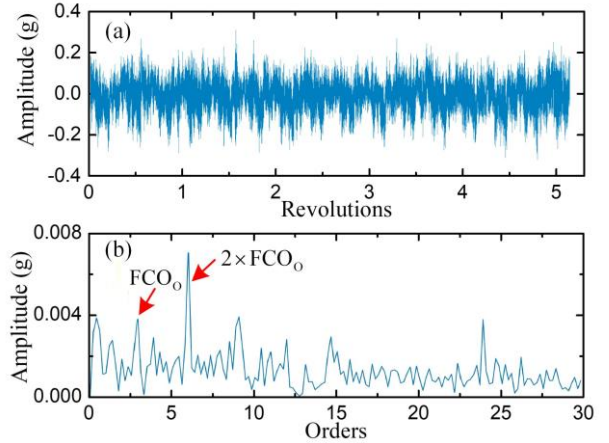
Fig. 6. (a) The resampled vibration signal at *x*-axis, and (b) envelope order spectrum.
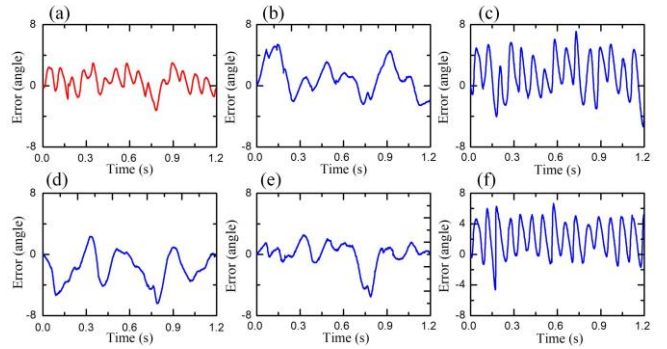
Fig. 7. (a) The angle error of the fused curve, (b)- (f) the angle errors of the selected 5 curves.

Finally, the rotating angle estimation accuracy is evaluated by comparing the curve $RA_f[n]$ and the measured

510

rotating curve obtained from the PMSM encoder. The results are shown in Fig. 7. The error of the fused curve is shown in Fig. 7(a), where it can be seen that the error fluctuates in the range of -4 degree to 4 degree. The errors of the 5 selected angle curves are shown in Figs. 7(b)-7(f). The error values are all larger than that of the fused curve. This result demonstrates that the fused curve has higher accuracy in rotating speed estimation, and hence benefiting diagnosis of bearing fault.

V. CONCLUSIONS

An improved rotating speed estimation algorithm is investigated for bearing fault diagnosis under the conditions of speed variation. The bearing vibration signals are acquired from a tri-axial accelerometer installed on the generator. The IFs of the vibration signals are extracted using the WSST algorithm. The rotating angle curves are computed from the IFs and then normalized according to the base rotating frequency. Total 9 normalized rotating angle curves are obtained from the three-channel of vibration signals and the correlation coefficients among the curves are computed. The curves with the highest correlation coefficients are selected and fused into a new curve. The fused curve has higher accuracy compared with the selected curves. The original bearing signal is resampled in accordance with the fused rotation curve estimated from the signal itself. The bearing fault type is determined from the spectrum of the demodulated resampled signal. The proposed method provides a high-accurate order analysis method under speed variation conditions without using a tachometer. The method can be used in rotation machine fault diagnosis under speed-varying conditions using only accelerometers.

REFERENCES

[1] S. Lu, R. Yan, Y. Liu, and Q. Wang, "Tacholess Speed Estimation in Order Tracking: A Review with Application to Rotating Machine Fault Diagnosis," *IEEE Trans. Instrum. Meas.,* vol. 68, pp. 2315-2332, Jul. 2019.

[2] Y. Qin, S. Xiang, Y. Chai, and H. Chen, "Macroscopic-microscopic attention in LSTM networks based on fusion features for gear remaining life prediction," *IEEE Trans. Ind. Electron.,* vol. 67, pp. 10865 - 10875, Dec. 2020.

[3] J. B. Xiong, Q. H. Zhang, G. X. Sun, X. T. Zhu, M. Liu, and Z. L. Li, "An Information Fusion Fault Diagnosis Method Based on Dimensionless Indicators With Static Discounting Factor and KNN," *IEEE Sens. J.,* vol. 16, pp. 2060-2069, Apr. 2016.

[4] C. Li, M. Liang, and T. Y. Wang, "Criterion fusion for spectral segmentation and its application to optimal demodulation of bearing vibration signals," *Mech. Syst. Signal Process.,* vol. 64-65, pp. 132-148, Dec 2015.

[5] Z. Y. Chen and W. H. Li, "Multisensor Feature Fusion for Bearing Fault Diagnosis Using Sparse Autoencoder and Deep Belief Network," *IEEE Trans. Instrum. Meas.,* vol. 66, pp. 1693-1702, Jul. 2017.

[6] J. W. Yang, Y. L. Bai, J. H. Wang, and Y. Zhao, "Tri-axial vibration information fusion model and its application to gear fault diagnosis in variable working conditions," *Meas. Sci. Technol.,* vol. 30, Sep 2019.

[7] S. Lu, X. Wang, Q. He, F. Liu, and Y. Liu, "Fault diagnosis of motor bearing with speed fluctuation via angular resampling of transient sound signals," *J. Sound Vib.,* vol. 385, pp. 16-32, Dec. 2016.

[8] J. Niu, S. Lu, Y. Liu, J. Zhao, and Q. Wang, "Intelligent Bearing Fault Diagnosis based on Tacholess Order Tracking for a Variable-Speed AC Electric Machine," *IEEE Sens. J.,* vol. 19, pp. 1850 - 1861, Mar. 2019.

Incremental Learning of Bearing Fault Diagnosis Via Style-Based Generative Adversarial Network

Yinjun Wang
College of Mechanical Engineering
Chongqing University
Chongqing 400044, China
20170701012@cqu.edu.cn

Liling Zeng
College of Mechanical Engineering
Chongqing University
Chongqing 400044, China
m2019zll@163.com

Xiaoxi Ding
College of Mechanical Engineering
Chongqing University
Chongqing 400044, China
dxxu@cqu.edu.cn

Liming Wang
State Key Laboratory of Mechanical Transmissions
Chongqing University
Chongqing 400044, China
lmwang@cqu.edu.cn

Yimin Shao*
State Key Laboratory of Mechanical Transmissions
Chongqing University
Chongqing 400044, China
ymshao@cqu.edu.cn

Abstract—At present, transfer learning of machine fault is a relatively popular research, its main problem is the imbalance of training data caused by the lack of actual fault data. The existing incremental learning model cannot solve the entanglement problem of sample features, and the ability to obtain new samples by combining features is limited. In this paper, Style-based Generative Adversarial Networks (StyleGAN) is used to map the data features to intermediate latent space, and then generate data by recombining features. StyleGAN realizes the complete separation of signal features. Therefore, StyleGAN can be used as a tool of data incremental learning to enrich the original data, solve the problem of imbalance between training data and test data, and achieve the goal of improving the accuracy of fault classification in the later stage. In the process of training, the category label is used as the auxiliary information to help the training model. The data of training set is enhanced, and the accuracy of fault diagnosis and classification is improved, the accuracy of fault classification network model is increased from 81.4% to more than 90%, so the validity of this method is proved.

Keywords—Data imbalance, incremental learning, StyleGAN, transfer learning, bearing fault diagnosis.

I. Introduction

In recent years, artificial intelligence has been widely used in the field of machine intelligent fault classification, and achieved good results. The traditional intelligent diagnosis methods assume that the training set and test set come from the same domain, which makes the input characteristics and data distribution characteristics the same. However, this assumption does not hold in some real-world machine learning scenarios. In some cases, training data of the model is difficult to collect [1]. For example, it is unlikely that the size and shape of the human fault are consistent with the size and shape of the actual mechanical faults, and we need to pay a high cost and a long time to obtain enough signals data in the actual fault. At present, the feasible method is that the training set of fault diagnosis network model is the laboratory data, or a large number of laboratory data and a small number of actual fault data as the training set. Both of these methods

will lead to the imbalance of the number of training sets and test sets in the network model of fault discrimination, so it is necessary to study the incremental learning of data.

At present, some scholars have discussed and studied the related problems of transfer learning and incremental learning. Shao *et al*. used auxiliary classifier Generative Adversarial Networks (GAN) to learn network learning signal features, and then generates one-dimensional signal data [2]. Wang *et al*. proposed a method of combining GAN and automatic encoder, which solves the problem of less samples to a certain extent [3]. Zhao *et al*. proposed a semi supervised model of adversarial generation network and used it to classify faults[4]. Aiming at the problem of insufficient training samples of rotating machinery, Ding *et al*. [5] proposed a model of GAN with small sample as input, and uses this structure to classify faults.

These incremental learning models do not solve the problem of feature entanglement, and its ability to get new samples by combining features is limited. StyleGAN maps the features of the data to the hidden vector and solves the problem of feature entanglement. StyleGAN can generate new data by combining different features [6]. The incremental learning methods are summarized as follows.

1) The fault vibration signal of bearing is collected, the fault signal is transformed into wavelet coefficient by wavelet transform, and then the wavelet coefficient is transformed into gray-scale image.
2) The resulting gray-scale image and corresponding labels are used as input of StyleGAN, and a new fault gray-scale time-frequency image is generated by the trained network.
3) The original and the generative gray-scale time-frequency graph are used as residual networks (Resnet) training set to classify bearing faults.

The rest of this paper is arranged as follows. Second II introduces the knowledge background of the StyleGAN. Section III describes the system architecture, model training methods. In the Section IV, we briefly introduce the incremental learning, and discuss its limitations. Section V provides the conclusion.

---

* Yimin Shao is the corresponding author. (e-mail: ymshao@cqu.edu.cn)

II. THEORETICAL BACKGROUND

A. Wavelet Transform

The dynamic characteristics of nonstationary vibration signal can be obtained by time-frequency analysis. The wavelet transform is used to process the original signal to get the wavelet coefficients, and then the appropriate size matrix is obtained by the down sampling. The wavelet transform uses a family of functions to represent signals, which are called wavelet function system. We suppose that function $\Psi \in L^2(R) \cap L^1(R)$, and $\Psi(0)=0$ is a family of functions obtained by stretching and translating $\Psi_{a,b}$:

$$\Psi_{a,b}(t) = |a|^{-1/2}\,\Psi\!\left(\frac{t-b}{a}\right), \quad a,b \in R, a \neq 0 \tag{1}$$

where $\{\Psi_{a,b}\}$ is wavelet, Ψ is basic wavelet or mother wavelet, a is the expansion factor, b is the translation factor. For signal $f \in L^2(R)$, its continuous wavelet transform is defined as follows.

$$W_f(a,b) = \langle f, \Psi_{a,b}\rangle = |a|^{-1/2}\int_{-\infty}^{+\infty} f(t)\overline{\Psi\!\left(\frac{t-b}{a}\right)}dt \tag{2}$$

where $\overline{\Psi}(t)$ is the complex conjugate of $\Psi(t)$, $\langle f, \Psi(t)_{a,b}\rangle$ represents the inner product of f, $\Psi_{a,b}$.

B. Style Generative Adversarial Networks

The StyleGAN consists of the discriminator D and generator G, shown in Fig. 1. These two parts are independent training. The generator starts from a random constant z which belongs to potential space $\mathbb{Z}$. The latent code z is normalized by pixel, the implementation method is as follows.

$$\overline{z}_i = z_i \,/\, \sqrt{\frac{1}{N}\sum_{i=1}^{N} z_i^2 + \varepsilon} \tag{3}$$

where z_i is the latent code, $N=512$, $\varepsilon = 10^{-8}$. The potential space $\mathbb{Z}$ is mapped to W by nonlinear mapping network. The W plays a role in controlling the style parameter $y=(y_s,y_b)$ of adaptive instance normalization(AdaIN). The convolution layer of synthesis network G is followed by a layer of AdaIN. The AdaIN s calculated as follows.

$$\text{AdaIN}(x_i, y) = y_{s,j}\frac{x_i - \mu(x_i)}{\sigma(x_i)} + y_{b,j} \tag{4}$$

where x_i is the feature map after being normalized separately, the style parameter $y=(y_s,y_b)$ is obtained by mapping network. The y_s is the scaling factor, the y_b is the offset factor. Therefore, the dimension of parameter y is twice that of feature maps. A constant tensor with a size of $4\times4\times512$ is used as the input of the synthesis network g, then add with noise. The output of the last layer of Style-based generator is converted to data by 1×1 convolution layer. The downscaling real data and generated data as input parameters of discriminator, the function of discriminator is to distinguish the source of data as much as possible.StyleGAN selects the entropy cross loss function as follows [7]:

$$L = \underbrace{\mathrm{E}_{\tilde{x}\sim P_{\tilde{x}}}[D(\tilde{x})] - \mathrm{E}_{x\sim P_x}[D(x)]}_{\text{Original critic loss}} + \lambda\,\underbrace{\mathrm{E}_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}}D(\hat{x})\|_2 - 1)^2]}_{\text{Gradient penalty}} \tag{5}$$

where $\tilde{x}=G(z)$ and according to $P_{\tilde{x}}$ distribution, original critic loss enables the discriminator to identify the source of the picture as much as possible.
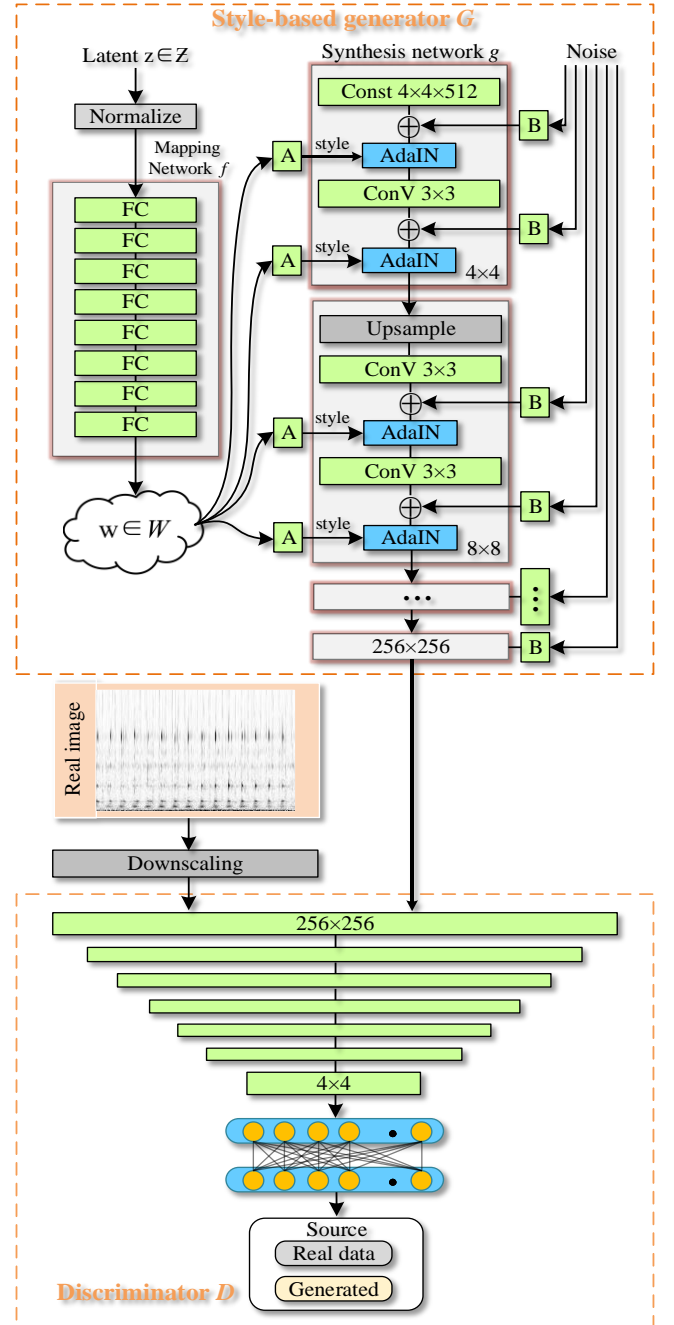
Fig. 1. Model of Style Generative Adversarial Networks

In Fig. 1, the Mapping Network f map $\mathbb{Z}$ to a latent space W, which controls the style of generator through Adai. Gaussian noise is added after each convolution layer. The A stands is an affine transform, which controls the style of image generation. The B is the scale factor to control the noise. The mapping network f consists of 8 full connection layers. The synthesis network g consists of 14 modules which consists of two convolution layers, two noise layers and two style control layers. The generated image and the reduced dimension image of the original image are used as the input of the discriminator. The discriminator consists of 7 convolution module layers and 2 full connection layers. The discriminator and generator are trained separately, the function of discriminator is to identify whether the input image is real or generated, the function of the generator is to generate images of different styles. After training to the best, they will reach Nash equilibrium.

C. Residual Networks

ResNet solves the problem of deep network gradient degradation, Compared with the traditional deep CNN, which has higher recognition accuracy and faster training speed [8]. The ResNet structure used in this article is shown in Fig. 2.
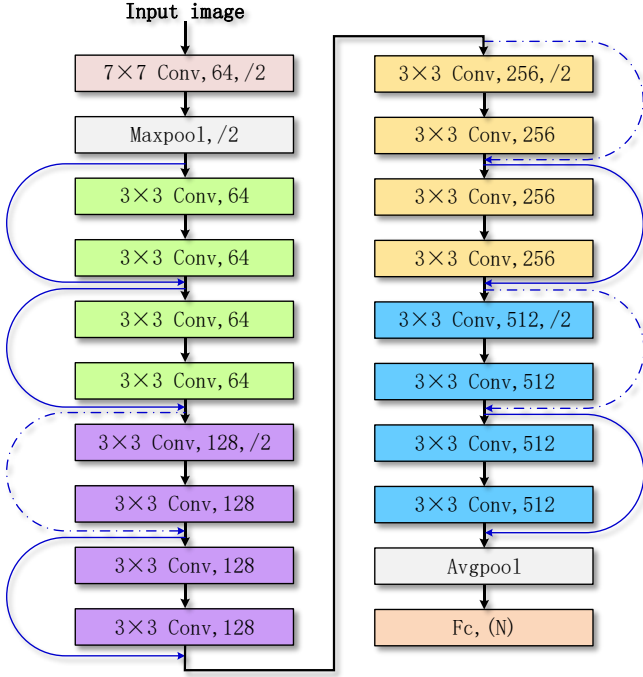
Fig. 2. Model of ResNet

Figure 2 is the architecture of 18-layers ResNet, where '64' means that the dimension of convolution kernels is 64, and '/2' indicates the movement of the convolutional kernels with a stride of 2. The dotted lines in the Fig. 2 indicate that the dimensions between modules are not the same, while the implementations represent the same. The operation of the dotted line is to expand the dimension of the input data by 1×1 convolution, then add the input and output of the module. The operation implemented is to directly add the input and output of the module, because the dimensions of the input and output are the same.

The residual learning module is shown in Fig. 3. When the input x and output y dimensions of the module are not equal, $y=\mathrm{F}(x)+W_s x$, that is the implementation process of the dotted line module in Fig. 2. When the input and output dimensions of the module are the same, then $y=\mathrm{F}(x)+x$, that is the implementation process of the solid line module. It can be found that the gradient of residual module is greater than 1, so no matter how many layers of residual modules are included in the residual network, its gradient will not disappear. ResNet solves the problem that the gradient of deep network disappears, which is the reason why we choose residual network as discrimination network.
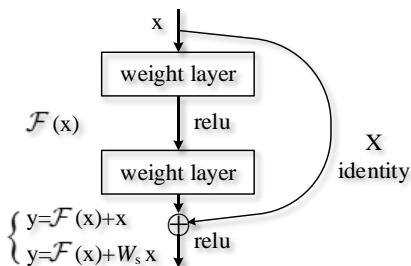
Fig. 3. Residual learning module.

III.   SYSTEM FRAMEWORK AND MODEL TRAINING

A. System Framework Design

In practical application, the fault identification accuracy of the model is generally not high using unbalanced data as training set. We build a framework using StyleGAN to generate high-quality artificial fault signals for data augmentation. The architecture consists of two parts, one is data generation based on StyleGAN, the other is fault classification via ResNet, as shown in Fig.4.
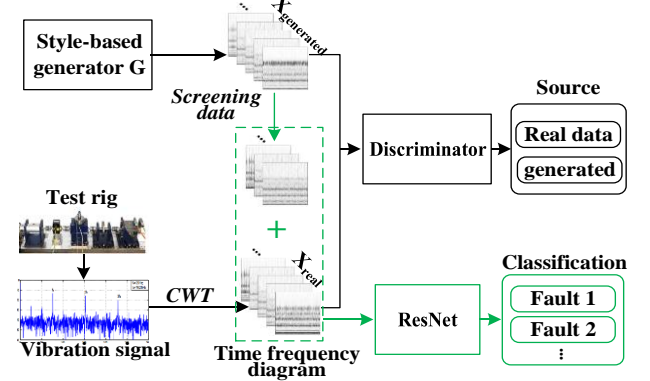
Fig. 4. Data augmentation strategy using StyleGAN for fault diagnosis task.

The structural details of Style-based generator G and Discriminator D are shown in Fig.1. The generator and discriminator optimized by (5). The goal of the generator is to make the generated data cheat the discriminator as much as possible. The object of the discriminator is to distinguish the source of the data as much as possible, that is, whether the data is real or generated. Only the signal data of a single fault is input each time, so the label of the generated data is also this fault type. The data of different faults are incremental learning by analogy. After samples generation, the generated data set is filtered, i.e. the data with wrong label is eliminated, and dataset has been augmented based on $X_{generated}$ after data filtering. Then we combine the high-quality manual fault data and the actual fault data as the input of ResNet. Through continuous iterative training of network parameters, the network can distinguish fault types more accurately.

We set up a test platform to collect vibration signals of different fault types, calculate the wavelet coefficient of vibration signal according to (2), and then through down sampling and gray-scale processing, the best time-frequency gray-white image of vibration signal is obtained. The calculation formula of transforming wavelet coefficients into gray-scale image is as follows.

$$x_{i,j}^{*} = \frac{x_{i,j} - X_{\min}}{X_{\max} - X_{\min}} \times 256 \qquad (6)$$

where $x_{i,j}$ is two dimensional wavelet coefficients, because the pixels of the image are 256×256, so the value range of i and j is 0-256. $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the wavelet coefficients of all the signals involved in the operation. $x^*_{i,j}$ is the normalized wavelet coefficient and the pixel value of the gray scale image. The value of wavelet coefficient is transformed into 0-256 by formula, and then the normalized wavelet coefficient matrix is transformed into 256×256 pixels gray-scale image.

We input the gray-scale images of different fault categories into StyleGAN to get the gray-scale images of corresponding categories, generating images and real

images as training set of ResNet. By iterating repeatedly, the objective function is minimized, and ResNet network parameter training completed.

B. Model Training Procedure

StyleGAN: The loss function of the network model is shown (5), and the optimal solution of its parameters is obtained through continuous iteration. The model training adopts Adam optimizer [18], the learning rate of network model is 0.001. In the period of StyleGAN training, the training process can be expressed as three steps.

1. The generator randomly generates time-frequency graph with noise.

2. Firstly, the discriminator is trained. The generated image and the real image are mixed together as the input of the discriminator. The goal of training is to minimize the loss function of the discriminator. In this process, the network parameters of the generator are fixed.

3. Then the generator is trained to generate images of different styles by randomly Z as input. The goal of the generator is to make the generated image as confusing as possible to the discriminator.

Continue to repeat steps 2 and 3 above until the total objective function reaches the minimum value or the required training times are completed. After many training iterations, the generator and discriminator are optimized continuously, so that the parameters are optimized and the loss function is minimized. After enough iterations, the generator and discriminator reach Nash equilibrium, so the generator can generate data similar to the real data, and the new data conforms to the distribution of the real data. Because StyleGAN has the ability to completely separate image features, the image with feature combination can be obtained by controlling the style of the generated image.

ResNet: The residual network takes the loss function (7) [8] as the objective function, uses Adam optimizer to train the model, the learning rate of network parameters is set to 1×10^{-5}. The partial differential of all self-illumination distribution in the objective function is calculated, multiplied by the learning rate, and then as the minuend number, the objective function is continuously optimized. Until the objective function reaches the minimum value or reaches the set number of cycles without change.

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{i,k} \log p_{i,k} \qquad (7)$$

IV. EXPERIMENTAL RESULTS

A. Description of Test Bench and Test Data

This paper adopted bearing data sets published by KAt DataCenter of Universität Paderborn, The paper [9] which published by Christian Lessmeier *et al.* have detailed introduction to the test device and measurement data.

The following is a brief introduction to the test equipment and measurement data. The test rig consists of the following main modules, (1)a drive motor, (2) atorque sensor, (3) a bearing test unit, (4) a flywheel and (5) a load motor, show in Fig. 5.
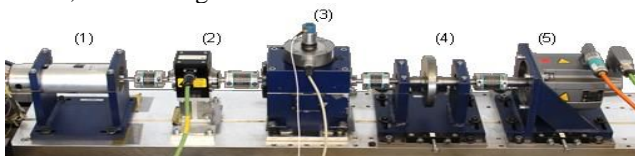
Fig. 5. Modular test rig.

The parameters of the test bearing are shown in Table I [9]. All three parameters were kept constant for the time of each measurement. At the basic setup (Set no. 0) of the operation parameters, the test rig runs at n = 1,500 rpm with a load torque of M = 0.7 Nm and a radial force on the bearing of F = 1,000 N. Three additional settings are used by reducing the parameters one by one to n = 900 rpm, M = 0.1 Nm and F = 400 N (set No. 1-3), respectively. For each of the settings, 20 measurements of 4 seconds each were recorded. Another parameter is the temperature, which was kept roughly at 45-50 °C during all experiments.

TABLE I. OPERATING PARAMETERS

No.	Rotational speed (rpm)	Load Torque (Nm)	Radial force (N)	Name of Setting
0	1500	0.7	1000	N15_M07_F10
1	900	0.7	1000	N09_M07_F10
2	1500	0.1	1000	N15_M01_F10
3	1500	0.7	400	N15_M07_F04

The test faulty bearing comes from two sources: one is produced by artificial damage, the other is produced by accelerated lifetime tests. There are three methods to make artificial faults, which are wire cut electrical discharge machining, rotary drilling and electrical discharge machining, as shown in Fig. 6. The bearing damaged by accelerated life test is shown in the Fig. 7. It can be found that the outer ring fault of EDM is similar to that of accelerated life damage.

Sharp trench by EDM - KA01   Drilling - KA09   Artificial pitting by electric engraver - KA03
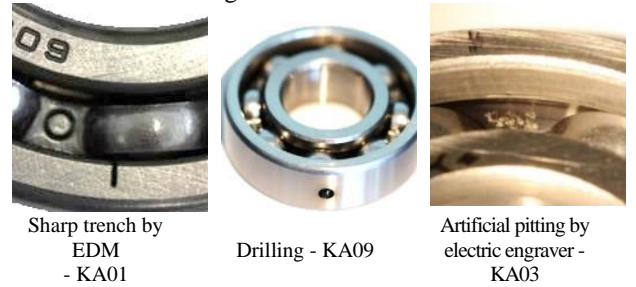
Fig. 6. Example drawing of bearing damaged artificially.
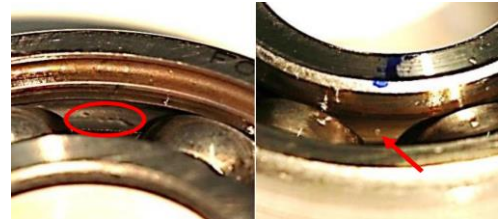
Fig. 7. Example of bearing damage in accelerated life test

TABLE II. OPERATING PARAMETER OF BEARINGS DURING RUN-IN PERIOD

Bearing Source	Artificial Damage Bearings		Accelerated Life Test Bearing		Healthy Bearings
Extent of Damage(level)	Inner Ring	Outer Ring	Inner Ring	Outer Ring	
1	KI01 KI03 KI05	KA01 KA05 KA07	KI04 KI14 KI17 KI21	KA04 KA15 KA22 KA30	K001 K002 K003 K004 K005 K006
2	KI07 KI08	KA03 KA06 KA08 KA09	KI18	KA16	

In this paper, 28 kinds of bearing failure model data are selected, including 12 kinds of artificial failure bearing, 10 kinds of accelerated life cycle failure bearing and 6 kinds of healthy bearing for comparison. They include bearings with level 1 and level 2 damage, and bearings with level 3

damage and composite damage are not selected. The selected test data are shown in Table II.

B. Data Processing

In the transfer learning of mechanical fault, the actual fault data is difficult to obtain. Most of the training data for fault diagnosis network are obtained in the laboratory, and the type of fault is also manufactured artificially. It is difficult to keep the size and shape of these faults from the actual fault size and shape. Therefore, this is the main reason for the low diagnostic accuracy of mechanical fault transfer learning. The training set of the test diagnosis network is artificial fault data, and the test set is accelerated life damage data.

There are 4 working conditions (as shown in Table I) for each fault size, and 20 groups of vibration data are collected for each working condition. We cut each group of data into 20 equal length data segments, and then get the time-frequency diagram of vibration signal through wavelet transform, down sampling and gray-scale processing. That is to say, each fault size can get 4×20×20=1600 real vibration time-frequency graphs. According to the working conditions in Table II, the number of corresponding time-frequency gray-scale images is obtained. Taking the corresponding time-frequency image as the real input of StyleGAN, the time-frequency image with corresponding labels is obtained, and the original data is incremental learning.

Fig. 8. Classification accuracy without incremental learning

C. Performance Comparisons

The accuracy of bearing fault classification without incremental learning is shown in Fig. 6. It can be found from this figure that the error classification occurs in the outer circle and inner circle faults, and the correct rate of normal bearing classification is 100%. Therefore, only the data of outer ring and inner ring need to be incremented.

Through the incremental learning of StyleGAN, we get the time-frequency gray-scale images of inner and outer rings, as shown in Fig. 9. It can be found that the generated gray-scale image has a high similarity with the actual gray-scale image, but it is not completely consistent. This is caused by the reason of adding noise and multi-feature fusion in the generated image which plays the purpose of enriching the original data.
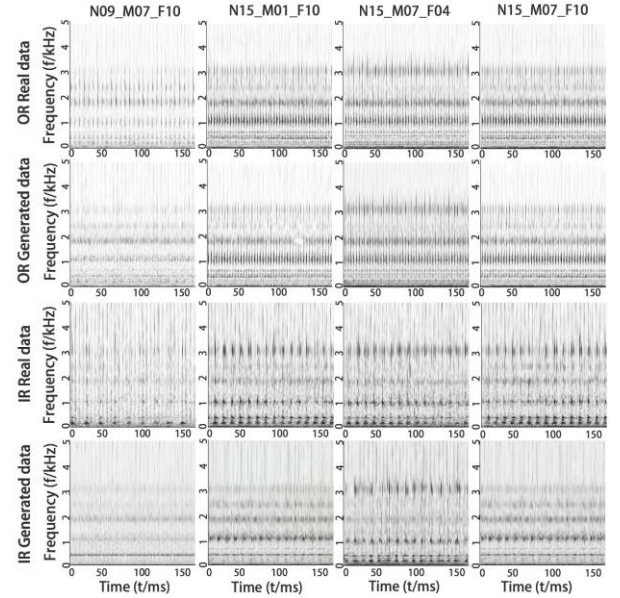
Fig. 9. Comparison between real time-frequency gray-scale image and generated time-frequency gray-scale image

By modifying the random seed and the function of style scale ψ, different time-frequency grayscale images can be generated by StyleGAN. Table III shows the combination of 12 training sets and the accuracy of their network training the classification effects of StyleGAN, ACCGAN and SNSSGAN incremental learning are compared the specific data are shown in Table III. It can be found that the accuracy of classification is improved by WGAN and ACGAN incremental learning, but it is not very obvious. After StyleGAN incremental learning, the accuracy of fault classification has been significantly improved. By observing A-B-C-D, we can see that the increase of the number of samples does not improve the accuracy of classification, which shows that 3000 training samples have made the network training saturated. The high fault rate of classification is not caused by the insufficient number of training samples, but by the single type of samples. It can be found that increasing the generated data can improve the StyleGAN accuracy of classification from the comparison

TABLE III. THE NUMBER AND ACCURACY OF TRAINING SAMPLES

Name		A	B	C	D	A′	B′	C′	D′	A″	B″	C″	D″
Real data	Normal	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	IR data	1000	1500	2000	2500	500	500	500	1000	0	0	0	0
	OR data	1000	1500	2000	2500	500	500	500	1000	0	0	0	0
Generated IR data		0	0	0	0	500	1000	1500	1500	1000	1500	2000	2500
Generated OR data		0	0	0	0	500	1000	1500	1500	1000	1500	2000	2500
Original accuracy		80.4%	81.3%	81.5%	81.4%	----	----	----	----	----	----	----	----
WGAN accuracy		----	----	----	----	81.7%	82.2%	83.1%	81.3%	82.5%	83.8%	83.9%	82.6%
ACGAN accuracy		----	----	----	----	81.9%	84.2%	84.5%	83.7%	82.5%	84.4%	86.7%	83.3%
StyleGAN accuracy		----	----	----	----	90.6%	91.7%	93.6%	91.9%	93.3%	94.3%	95.8%	93.9%

of A-A '-A", B-B '- B", C-C '- C", D-D '- D' 'in Table III. It proves the high quality of the generated data, enrich the original data set and improve the accuracy of fault classification.

V. CONCLUSION

In this paper, we propose an incremental learning framework based on StyleGAN, which can separate the features of signal data unsupervised. It controls different levels of features through different layers of style. It realizes the mixing and recombination of the features of the original signal data and synthesizes the generated data with various features. Therefore, the StyleGAN incremental learning framework achieves real data expansion, instead of simply copying the original data. In fault transfer learning, the data tested in the laboratory is quite different from the fault data actually generated. This leads to the problem of low diagnostic accuracy in fault transfer learning, which is well solved by StyleGAN incremental learning framework. The classification results show that the model can generate convincing data and can be used as a data enhancement technology in transfer learning. The time-frequency graph is generated by the random combination of the features, which cannot be accurately controlled to get the desired graph, which needs further research.

VI. ACKNOWLEDGEMENT

REFERENCES

[1] S. Pan, and Q. Yang, "A Survey on Transfer Learning. "*IEEE Transactions on Knowledge & Data Engineering*, vol. 22, DOI 10.1109 -/TKDE. 2009.191, no. 10, pp. 1345-135, 2010.

[2] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis."*Computers in Industry*, vol. 106, pp. 85-93, 2009.

[3] Z. Wang, J. Wang, and Y. Wang, "An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition," *Neuro- computing*, vol. 310, DOI 10.1016/j.neucom. 2018.05.024, pp. 213-222, Oct. 2018.

[4] D. Zhao, F. Liu, and M. He, "Bearing Fault Diagnosis Based on the Switchable Normalization SSGAN with 1-D Representation of Vibration Signals as Input," *Sensors*, vol. 19, DOI 10.3390/s19092000, no. 9, Apr. 2019.

[5] Y. Ding, L. Ma, J. Ma, C. Wang, and C. Lu, "A Generative Adversarial Network-Based Intelligent Fault Diagnosis Method for Rotating Machinery Under Small Sample Size Conditions,"*IEEE Access*, vol. 7, DOI 10.1109/access.2019.2947194, pp. 149736-149749, Oct. 2019.

[6] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *computer vision and pattern recognition*, DOI 10.1109/CVPR.2019.00453, pp. 4401-4410, 2019.

[7] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville. "Improved training of Wasserstein GANs," 2017.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition, DOI 10.1109/CVPR.2016.90, 2015.

[9] Christian Lessmeier，James Kuria Kimotho, Detmar Zimmer and Walter Sextro, "Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification," *European Conference of The Prognostics And Health Management Society*, 2016

Recovery of Under-sampled Signal During High-speed Machining Condition Monitoring Using Approximate Sparsity in Frequency Domain

Yang Li
Shool of Aerospace Engineering
Xiamen University
Xiamen, China
liyangxdjx@163.com

Binqiang Chen[*]
Shool of Aerospace Engineering
Xiamen University
Xiamen, China
cbq@xmu.edu.cn

Nianyin Zeng
Shool of Aerospace Engineering
Xiamen University
Xiamen, China
zny@xmu.edu.cn

Xincheng Cao
Shool of Aerospace Engineering
Xiamen University
Xiamen, China
XCCAO12@163.com

Bin Yao
Shool of Aerospace Engineering
Xiamen University
Xiamen, China
yaobin@xmu.edu.cn

Abstract—**Aiming at the problem of under-sampled cutting force signal caused by unreasonable setting of sampling parameters in the high-speed machining condition monitoring system, a method of spectrum sensing based on the principle of approximate sparseness in frequency domain has been proposed. The non-linearity of machining system and sampling process makes the output signal of monitoring system contain higher harmonics, which shows obvious approximate sparsity on the Fourier basis. Using frequency points with large peaks can achieve sparse approximation of the spectrum, and obtain several frequency subsets as the result. The principle of spectrum aliasing is used to calculate the actual frequency range of each frequency subset and correct Fast Fourier Transform (FFT) spectrum of the cutting force measurement. Experiments on high-speed milling of aluminum alloy verify the effectiveness of this method. The results show that the proposed method is effective enough to recover the actual waveform of the cutting force signal, and the relative envelope error between the recovery temporal wave and the test signal is less than 4%. The research results provide some engineering and technical support for applying sparse theory to recover under-sampled signals.**

Keywords—high-speed machining, under-sampled signal, Fourier basis, approximate sparse

I. INTRODUCTION

In recent decades, high-speed machining, with the advantages of high efficiency, high quality and low consumption, has gradually become a research hotspot in academia and industry community, and is widely used in the field of modern metal processing [1-4]. However, the milling system works in an extremely high-speed environment, which is easy to bring a decline in processing quality of parts as a result of tool wear. Even, it may damage the machine and endanger the life of workers in some serious cases. Therefore, it is of great significance to monitor the state of high-speed cutting systems and select appropriate monitoring signals. As ideal monitoring signals, vibration signal, acoustic emission and cutting force have been widely used in dynamics testing and analysis of high-speed cutting systems [5-7]. Among them, cutting force is favored as the signal which can effectively reflect the cutting process [8-9].

The classic data acquisition system is shown in Fig. 1.

Cutting force is measured by dynamometer and converted into an analog voltage signal f_a by an internal piezoelectric sensor and a charge amplifier. After low-pass filtering, the analog voltage signal is recorded in a computer by an ADC acquisition card for processing and analysis. In order to keep the effective information of cutting force completely, Nyquist–Shannon sampling theorem requires that the sampling frequency must be higher than 2 times the highest frequency of the analyzed signal [10]. Due to the nonlinearity of cutting system and sampling process, the milling force often contains high-order harmonics of spindle frequency. In the actual milling force experiment, when the sampling parameters are set unreasonably, the harmonic components exceeding the Nyquist frequency will be folded into the Nyquist zone through the filter steepness band of the anti-aliasing filter, which may give high-speed cutting system status analysis caused difficulties. In order to make full use of the existing measurements, it is important to develop an algorithm to recover the real signal from the under-sampled signal. Reference [11] proposed frequency estimation without fuzzy based on the time delay between sampling channels, and [12] chose an appropriate delay time to solve the problem of frequency aliasing under low-rate ADC. The idea of solving system of linear congruence equations by the Chinese remainder theorem provides a new way for the problem of spectrum estimation of under sampled signals. A variety of frequency estimation algorithms and improved algorithms based on the Chinese remainder theorem are proposed [13-15]. Although the above reconstruction algorithms can solve the problem of frequency aliasing on the degree, but it requires multiple ADCs, which cost a lot in practical application and still have certain limitations. HU et al. [16] reconstructed the blade vibration signal from the under-sampled signal based on the sparse characteristic reconstruction of signal, and provided a new idea for the single-channel reconstruction algorithm of the under-sampled signal.

Sparse representation was reintroduced by Donoho in 2001 [17]. Katabi et al. [18, 19] first proposed a kind of Sparse Fast Fourier Transform (SFFT) which based on the approximate sparsity of signal spectrum. H. hassanieh [20] applied the algorithm to the spectrum sensing, and realized the better recovery of the original signal spectrum with partial samples. Based on orthogonal polynomials to solve

---

*Binqiang Chen is the corresponding author. (e-mail: cbq@xmu.edu.cn).

the positions of K non-zero terms, Hsieh S constructed an under-sampled signal recovery algorithm. The experimental results show that the algorithm has achieved significant results in reducing complexity [21].
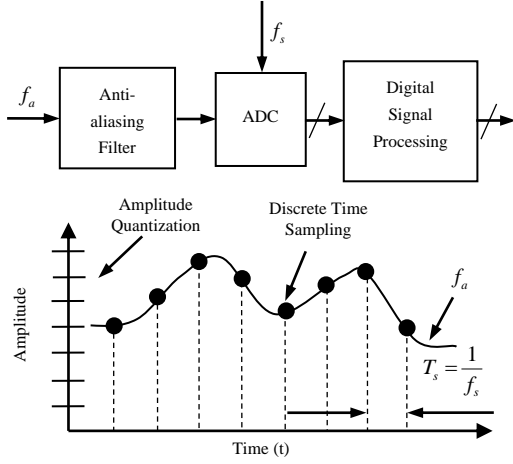
Fig. 1. Typical sampled data system.

Based on this, an under-sampled signal recovery method based on approximate sparsity in frequency domain is proposed for under-sampling feature of the condition monitoring signal during high-speed machining .This method achieves sparse approximation of the spectrum based on the approximate sparse characteristics of the cutting force signal in the frequency domain, observes the peak corresponding frequency points and divides the spectrum into several spectrum subsets, and calculates actual frequency of each subset by the principle of spectrum aliasing. The actual frequency of the subsets is used to realize spectrum correcting, and the Inverse Fast Fourier Transform(IFFT) is used to recover the signal's waveform in time domain. The results show that this method can effectively recover under-sampled signals and does not require the absolute sparsity of the test signal.

II. METHOD

The approach in our paper consists of three parts: approximate sparse theory in frequency domain, the principle of under-sampling spectrum aliasing, recovery algorithm. The architecture of the algorithm is plotted in Fig. 2.
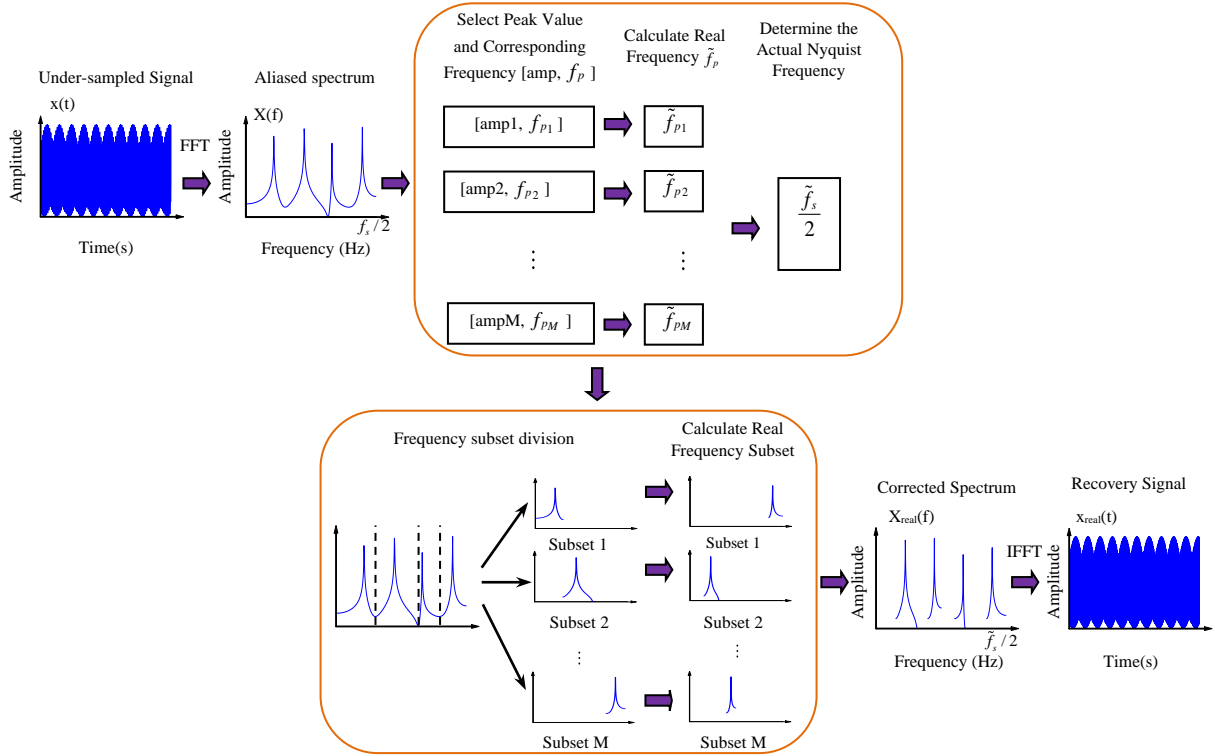
Fig. 2. The architecture of the algorithm.

A. Approximate Sparse Theory in Frequency Domain

Sparsity reflects the degree of energy concentration of the signal or in a certain transform domain. In general, the signal is not sparse. However, an appropriate orthogonal basis (Fourier basis, wavelet basis, Discrete Cosine basis, etc.) can be selected to describe the signal in the transform domain. As a result, the energy dispersive distribution in the spatial domain is transformed into the relatively concentrated energy distribution in the transform domain [22]. The discrete signal of length N can be expressed as a linear combination of orthogonal basis functions:

$$x = \sum_{i=1}^{N} \Psi_i c_i , \qquad (1)$$

where Ψ_i is the basis function, c_i is the representation coefficient. If there are only $K(K \geq N)$ non-zero entries in the n-dimensional vector $C=\{c_i\}_{i=1}^{N}$, C is a sparse vector whose sparsity is K, and x is called a K-sparse signal in the Ψ domain. In fact, strictly sparse signals are almost

non-existent. As long as the signal x of length N satisfies the conditions that there are $K(K \geq N)$ large nonzero value in C, x can be considered as a sparse signal approximately, which can be approximated by using K-sparse signal.

In digital signal processing, Discrete Fourier Transform (DFT) is known as one of the most basic and important signal analysis methods. DFT uses the basis function $e^{j2\pi kn/N}(k, n = 0,1,2,\cdots N-1)$ to represent the signal in the frequency domain to obtain the frequency characteristic. However, in practical engineering applications, background noise and the energy leakage caused by signal truncation both make representation coefficients contain a large number of redundant small coefficients as well as $K(K \geq N)$ large coefficients in frequency domain. According to the above-mentioned approximate sparse theory, the process of the approximate sparse decomposition and reconstruction on basis functions of DFT can be obtained, as shown in Fig. 3.
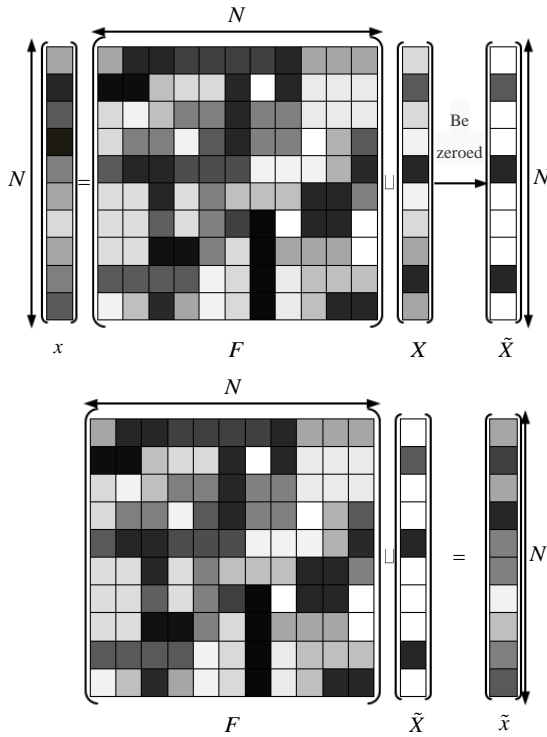
Fig. 3. The approximate sparse decomposition and reconstruction model in the frequency domain.

As describe above in Fig. 3, x is a digital signal obtained from sampled data system. It can be seen that there are lots of small value (light grid) in coefficient vector X got by decomposing x on DFT basis. The approximately sparse method for x in frequency domain is zeroing the small coefficients in X to obtain the vector $\tilde{X}$. As a result, approximate signal $\tilde{x}$ is acquired by using Inverse Discrete Fourier Transform (IDFT) on $\tilde{X}$.

B. The Principle of Undersampling Spectrum Aliasing

Under certain conditions, an analog signal can be completely represented by the value at the sampling points of equal time intervals, and these sample values can be used to recover the signal. This property comes from Nyquist–Shannon sampling theorem (hereinafter referred to as sampling theorem). The sampling theorem states that the sampling frequency must not be less than twice of the highest frequency in the signal:

$$2f_H \text{ £ } f_s \qquad (2)$$

where f_s is sampling frequency, f_H is the highest frequency in the signal. Otherwise, the signal will be under-sampled, resulting in aliasing in the frequency domain, as shown in Fig. 4.

In Fig. 4, the first Nyquist zone is defined to be the frequency range from 0Hz to $f_s/2$. The frequency spectrum is divided into an infinite number of Nyquist zones, each having a width equal to $f_s/2$ as shown. In practice, the ideal sampler is replaced by an ADC followed by an FFT processor. The FFT processor only provides an output from 0Hz to $f_s/2$, i.e., all the signals of interest lie in the first Nyquist zone. Fig.4 (a) shows the band of sampled signals is located in the first Nyquist zone, so the frequency band can be output correctly. Fig.4 (b) shows that the band of sampled signals is located in the second Nyquist zone, and the output frequency band is the image result of the original frequency band. Fig.4 (c) shows that the band of sampled signals is located in the third Nyquist zone, and the output frequency band is the secondary image result of the original frequency band. Therefore, when the sampling frequency of the signal is lower than the Nyquist frequency, the spectrum obtained by Discrete Fourier Transform is the result of multiple mirroring. The relationship between it and the spectrum of the input signal can be described by (3), which is called the principle of under-sampling spectrum aliasing.
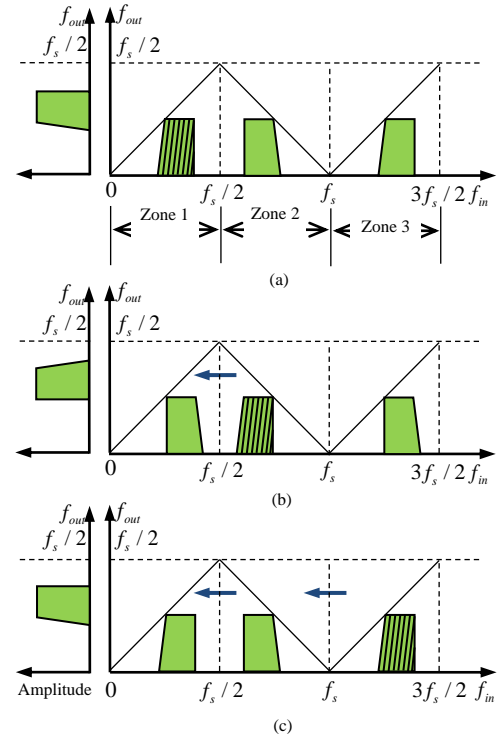
Fig. 4. Undersampling and frequency translation between Nyquist eones

$$X(f_{out}) = \begin{cases} X(f_{in}) & \overset{\underset{\mathrm{}}{}}{=}\min_{k=0}^{\infty}|kf_s - f_{in}|, \ kf_s < f_{in} \\ conj\{X(f_{in})\} \overset{\underset{\mathrm{}}{}}{=}\min_{k=0}^{\infty}|kf_s - f_{in}|, \ kf_s > f_{in} \end{cases} \qquad (3)$$

C. Recovery Algorithm

- Discrete time sequence of under-sampled signal during high-speed machining condition monitoring is $\{x(n)|n=1,2,K,N\}$, $N \in \mathbb{N}^{+}$. The sampling frequency is f_s ,and the spindle rotating frequency is f_a . Procedures of the recovery algorithm based on approximate sparse theory in spectral domain are as follows:

- Perform FFT on the cutting force sequence.

$$X(f)=FFT\{x(n)\} \tag{4}$$

- Mark M spectral bins with larger amplitude and calculate the true frequency $\{\tilde{f}_{pi}\}_{i=1}^{M}$ of each bin.

$$\tilde{f}_{pi}=\begin{cases} kf_s+f_{pi} \text{ , while } (kf_s+f_{pi})/f_a \in \mathbb{N}. \\ kf_s-f_{pi} \text{ ,while } (kf_s-f_{pi})/f_a \in \mathbb{N}. \end{cases} \tag{5}$$

- Segment the first Nyquist zone in the range of $[0, f_s / 2]$ into M frequency subsets.

$$X(f)=\begin{cases} X_1(f) & f \in (f_{1,l},f_{1,h}) \\ X_2(f) & f \in (f_{2,l},f_{2,h}) \\ \quad\vdots \\ X_M(f) & f \in (f_{3,l},f_{3,h}) \end{cases} \tag{6}$$

where $f_{1,l}=0$, $f_{i,l} \neq f_{i,h}$, $f_{i,h}=f_{(i+1),l}$,$f_{M,h}=f_s / 2$.

- Calculate the actual frequency range of $X_i(f)$.

$$\begin{cases} \tilde{X}_i(kf_s+f)=X_i(f) & \text{while } (kf_s+f_{pi})/f_a \in \mathbb{N} \\ \tilde{X}_i(kf_s-f)=conj\{X_i(f)\} & \text{while } (kf_s-f_{pi})/f_a \in \mathbb{N} \end{cases} \tag{7}$$

- Construct consequence $\{X_{real}(f)|0 \leq f \leq \tilde{f}_s/ 2\}$ with $\tilde{f}_s / 2=[2max(\{\tilde{f}_{pi}\}_{i=1}^{M}) / f_s]f_s / 2$. $[\ ]$ is integer-valued operator.

$$\tilde{X}_{real}(f)=\{\tilde{X}_1(f),\tilde{X}_2(f),L ,\tilde{X}_M(f)\} \tag{8}$$

Then, the actual spectrum of cutting force is expressed as

$$X_{real}(f)=\begin{cases} \tilde{X}_{real}(f) & 0 \leq f \leq \tilde{f}_s/ 2 \\ conj\{\tilde{X}_{real}(\tilde{f}_s- f)\} & \tilde{f}_s/ 2 \leq f \leq \tilde{f}_s \end{cases} \tag{9}$$

- Apply IFFT on the actual spectrum $\tilde{X}_{real}(f)$ to obtain the recovery signal.

$$x_{real}(n) = IFFT\{X_{real}(f)\} \tag{10}$$

III. EXPERIMENTS

The cutting experiments were conducted on a milling center of the type Mazak FJV-200 UHS. The work piece was a7505 aluminum alloy solid block of size that is bolted to the Kistler 9265B dynamometer. The spindle rotation frequency is 292Hz. During the collections of the cutting force signal, the sampling frequency was set as 2000Hz and the anti-aliasing filtering operation was not activated. The temporal waveform and the FFT spectrum of the cutting force measurement of a certain channel are shown in Fig. 5.
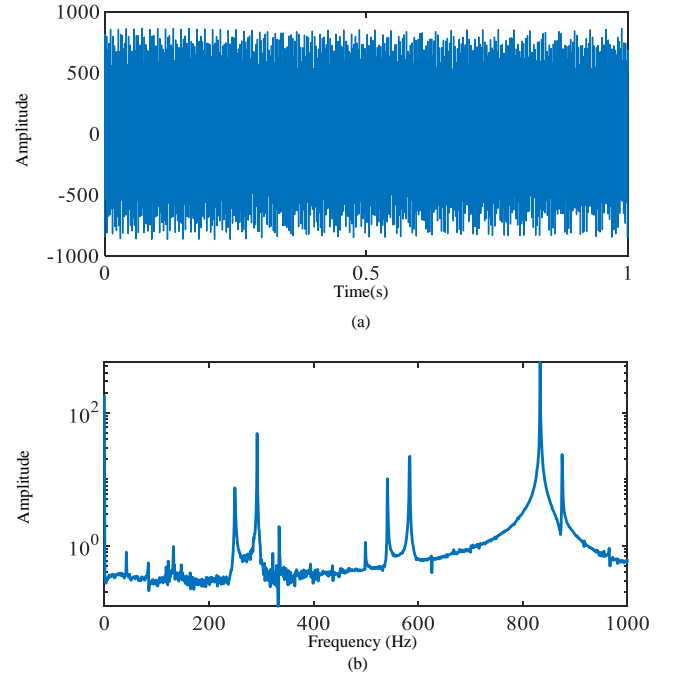
Fig. 5. (a)Temporal waveform of the cutting force measurement; and (b) the FFT spectrum of the cutting force measurement.

Since the anti-aliasing filter is not applied, higher harmonics of the spindle rotation frequency are aliased to the first Nyquist zone. As such, ten peak points and their corresponding frequency points are recognized in Fig. 5 (b). It can be calculated that the true frequency of each frequency point based on the principle of under-sampling spectrum aliasing. According to the index of frequency point, the first Nyquist zone can be divided into ten frequency subsets, as shown in Fig. 6. It is observed that the temporal waveform of each frequency subset can be approximated by sinusoid or cosine. The true ranges of each frequency subset calculated refer to(7) are listed in Table 1.

TABLE I.   FREQUENCY COMPARISON

No.	Frequency Range and Nyquist Zone			
	Frequency Range in the Spectrum (Hz)	Nyquist zone in the Spectrum	Actual Frequency Range (Hz)	Actual Nyquist zone
1	[0,64]	1	[2000,2046]	3
2	[65,142]	1	[6065,6142]	7
3	[143,270]	1	[1730,1857]	2
4	[271,313]	1	[271,313]	1
5	[314,350]	1	[2314,2350]	3
6	[351,520]	1	[3480,3649]	4
7	[521,560]	1	[1440,1479]	2
8	[561,625]	1	[561,625]	1
9	[626,870]	1	[1130,1374]	2
10	[871,1000]	1	[871,1000]	1
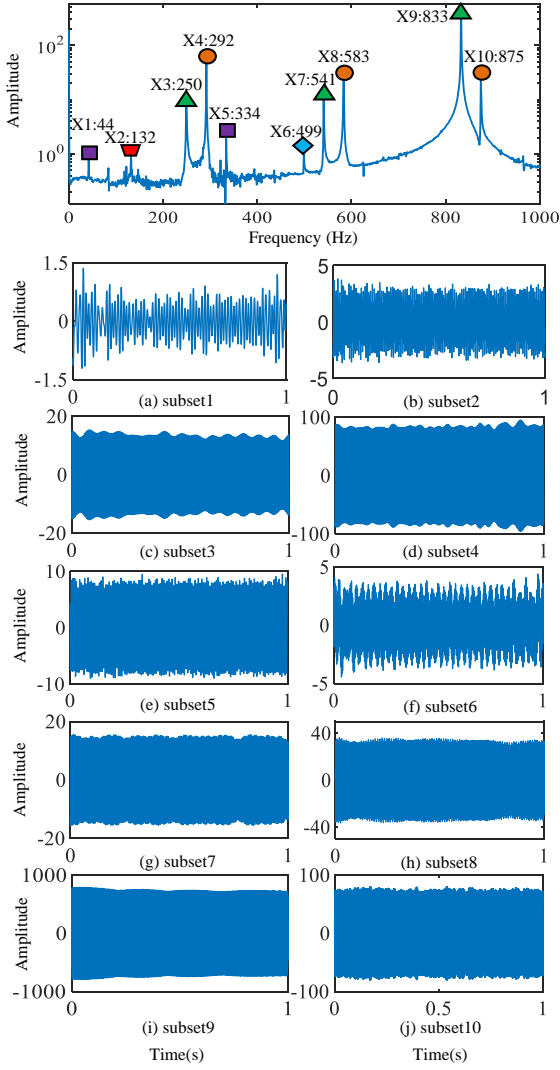
Fig. 6. Temporal waveform of cutting force frequency subsets.

(a)

(b)
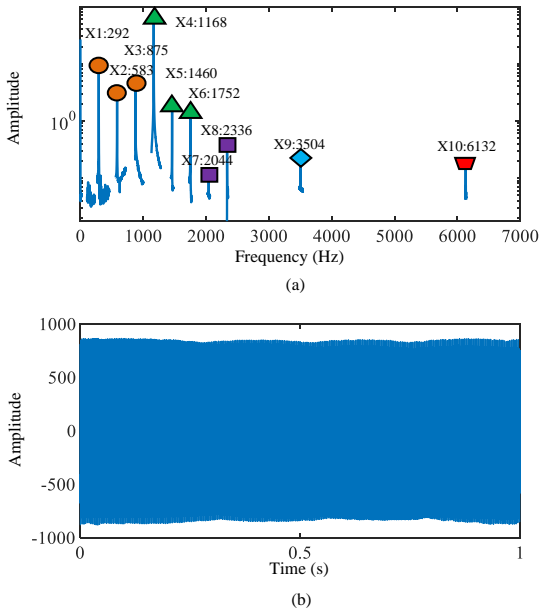
Fig. 7. (a) Corrected FFT spectrum of the recovery signal; and (b) the temporal waveform of the recovery signal.

Three frequency subsets (No.4,8,10) are normally sampled, while the others (No.1,2,3,5,6,7,9) are aliased. Since the actual highest frequency is 6142Hz, the first Nyquist frequency zone of the actual spectrum in Fig. 7 (a) is refactored to [0, 7000] Hz. Response exists at the rotation frequency and its frequency multiplier in the corrected spectrum. The recovery signal based on the corrected spectrum is plotted in Fig. 7 (b). The relative envelope error between the original measurement and the recovery signal is computed in Fig. 8. The relative envelope error E_R is less than 4%. It indicates a high recovery precision. Moreover, no end sample distortions are found, which means that the method proposed in this paper has high reliability and application value in the actual application.

$$E_R = \frac{envelope(x_{real}(n)) - envelope(x(n))}{envelope(x(n))} \times 100\% \quad (11)$$

where, $envelope(\cdot)$ is envelope function.
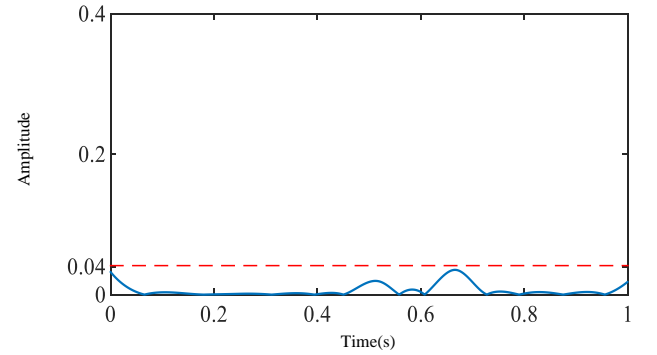
Fig. 8. The relative envelope error between the original measurement and the recovery signal
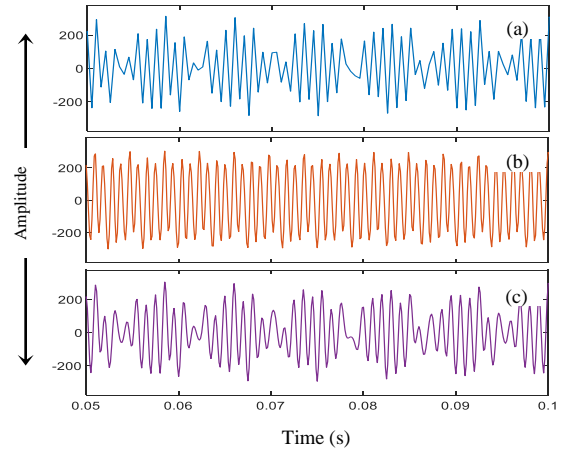
Fig. 9. Zoom-in plots of (a) the original measurement, (b) the econstructed measurement using the proposed method, and (c) interpolated easurement.

For comparison, zoom-in plots of the original measurement and the reconstructed measurement are also shown in Fig. 9(a, b). These two measurements are very different in the waveforms. In Fig. 9(c), the interpolated measurement based on the traditional cubic spline interpolation is plotted. Although the interpolated measurement offers denser and smother sampling, as a

construction method of approximation function in time domain, it depends on the original measurement highly. When the original measurement undersampled are approached by cubic spline function, the function must follow the change trend (the main performance of different harmonic components in time domain) of original measurement without adding new physical information, that is, the position of effective spectrum l lines in frequency domain does not change. The proposed method was to correct the original measurement in the frequency domain (mainly the position of the effective spectrum l line), that is to rectify the internal harmonic component information, which will inevitably change the trend of the time-domain waveform. These results have verified that the proposed method can overcome the deficiencies of traditional interpolation techniques.

IV. CONCLUSION

In this paper, we propose a new recovery method for under-sampled cutting force signal measured from high-speed machining condition monitoring. This method is based on the approximate sparsity of cutting force signal in frequency domain under the condition of constant speed. As a result, we can realize sparse approximation of test signal by sparse signal refer to approximate sparse theory in frequency domain. After FFT is used for the test signal, the spectrum sequence is divided into several continuous frequency subsets according to the index value of the frequency point with larger peak value. Combined with the principle of spectrum aliasing, the actual frequency range of subset is calculated and the corrected spectrum is obtained. Finally, IFFT is applied to the corrected spectrum to get the recovered signal. The experimental results show that the method has a good recovery result, and the relative envelope error between the recovery signal and the measured signal is less than 4%. This method can also provide a new idea for the analysis of under-sampled signals.

REFERENCES

[1] J. Q. Dang, H. Zhang, W. W. Ming, Q. L. An, and M. Chen, "New observations on wear characteristics of solid Al2O3/Si3N4 ceramic tool in high speed milling of additive manufactured Ti6Al4V," Ceramics International, 2020, pp. 5876-5886.

[2] F. Molaiekiya, M. Aramesh, and S. C. Veldhuis, "Chip formation and tribological behavior in high-speed milling of IN718 with ceramic tools," Wear, 2020, p. 203191.

[3] Z. Wang, and M. Rahman, "11.10 - High-Speed Machining," Comprehensive Materials Processing, 2014, pp.221-253.

[4] H. Zhang, J. Q. Dang, W. W. Ming, M. Chen, and Q. L. An, "Cutting responses of additive manufactured Ti6Al4V with solid ceramic tool under dry high-speed milling processes," Ceramics International, 2020.

[5] H. AlanA, "Study on Detection of Cutting Tool Edge Position by Acoustic Emission Technique," The Proceedings of Mechanical Engineering Congress, 2016.

[6] K. Jemielniak, and P. J. Arrazola, "Application of AE and cutting force signals in tool condition monitoring in micro-milling," CIRP Journal of Manufacturing Science and Technology, 2008, pp. 97-102.

[7] J. Ratava, M. Lohtander, and J. Varis, "Tool condition monitoring in interrupted cutting with acceleration sensors," Robotics and Computer-Integrated Manufacturing, 2017, pp. 70-75..

[8] J. Qiu, "Research on cylindrical turning process stability judgment based on dynamic cutting process," Journal of Mechanical Engineering, 2019, pp. 221-231.

[9] B. Zhang, C. L. Zhao, and B. C. Wen, "Dynamic cutting force measurement test and prediction of time series model for machine tools," Journal of Northeastern University(Natural Science), 2019, pp. 521-525+530.

[10] W. Kester, "MT-002 TUTORIAL What the Nyquist Criterion Means to Your Sampled Data System Design," Processcontrol.analog.com, 2007, pp.5–16.

[11] M. D. Zoltowski, and M. P. Mathews, "Real-time frequency and 2-D angle estimation with sub-Nyquist spatio-temporal sampling," Signal Processing IEEE Transactions on, 1994, pp.2781-2794.

[12] D. W. Tufts, and H. Ge, "Digital estimation of frequencies of sinusoids from wide-band under-sampled data," International Conference on Acoustics. IEEE, 1995.

[13] M. Saab, Y. Nasser, and K. Y. Kabalan, "Frequency Estimation of Multiple Components using Chinese Remainder Theorem," International Conference on Telecommunications. 2018.

[14] H. Sun, L. W. Chen, H. J. Zhang, S. Huang, and L. Yu, "Frequency estimation of multiple sinusoids with three sub-Nyquist channels," Signal Processing, 2017,139.

[15] H. Liang, Q. Zhang, and C. S. Yang, "A generalized robust Chiese Remainer Theorem and its application to frequency estimation with undersampling," Journal of Electronics and Information Technology, 2010, 32(8).

[16] J. Lin, Z. Hu, Z. S. Chen, Y. M. Yang, and H. L. Xu, "Sparse reconstruction of blade tip-timing signals for multi-mode blade vibration monitoring," Mechanical Systems & Signal Processing, 2016, pp.250-258.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," Siam Review, 2001, pp.129-159.

[18] H. Hassanieh, P. Indyk, D. Katabi, and E. Price, "Simple and Practical Algorithm for Sparse Fourier Transform," Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms(Proceedings: Society for Industrial and Applied Mathematics, 2012, pp.1183-1194.

[19] P. Indyk, M. Kapralov, and E. Price, "(Nearly) Sample-Optimal Sparse Fourier Transform," Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms(Proceedings: Society for Industrial and Applied Mathematics, 2013, pp.480-499.

[20] H. Hassanieh, L. Shi, O. Abari, E. Hamed, D. Katabi, "GHz-wide sensing and decoding using the sparse Fourier transform," IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, 2014, pp.2256-2264.

[21] S. Hsieh, C. Lu, and S. Pei, "Sparse Fast Fourier Transform by downsampling," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp.5637-5641.

[22] N. F. Dong, "Compressive sampling and reconstruction of sparse wideband signals," Nanjing University of science & Technology, 2019.

A Novel Weak Fault Diagnosis Method Based on Sparse Representation and Empirical Wavelet Transform for Rolling Bearing

Wei Lu
School of mechanical and electrical engineering
Beijing University of Chemical Technology
Beijing, China
m17761247385@163.com

Liuyang Song
School of mechanical and electrical engineering
Beijing University of Chemical Technology
Beijing, China
xq_0703@163.com

Lingli Cui
Laboratory of Advanced Manufacturing Technology
Beijing University of Technology
Beijing, China
acuilingli@163.com

Huaqing Wang*
School of mechanical and electrical engineering
Beijing University of Chemical Technology
Beijing, China
hqwang@mail.buct.edu.cn

Abstract—**Rotating machinery is widely used in industry. However, it works in tough environment, which makes the fault features extraction difficult. In the last few years, sparse representation, as a kind of effective feature extraction method, has great promise in industrial diagnosis. As for the traditional sparse representation, the greedy algorithm used to update the sparse coefficients is prone to produce local optimal solution and lead to over-fitting. In addition, if the signal contains a lot of redundant information, the basis learned by the traditional method cannot well represent the fault signal. Aim at the above questions, a novel weak fault diagnosis method based on sparse representation and empirical wavelet transform (EWT) is proposed in this paper. Gaussian filter is exploited to process the signal spectral, which can make the signal spectral smooth and the spectrum division more precise. Next, the signal spectrum is divided into N parts based on EWT. Then the kurtosis is utilized to screen out the optimal part of spectrum, which will be exploited to obtain a sparse basis. The constraint with nuclear norm is applied to remove the redundant component of the basis. Finally, the *LASSO* with elastic net, is employed to get sparse signal, the envelope spectrum is used to extract fault feature. Experimental results show that this method is better than traditional sparse representation using learning dictionary.**

Keywords-Sparse representation; Empirical wavelet transform; LASSO with elastic net; Gaussian blur; nuclear norm.

I. INTRODUCTION

Rolling bearing is critical for rotating machinery, and 40% of machinery faults suffer from bearing damage. So the fault detection, isolation and identification of a rotating bearing are significant [1,2]. In recent years, sparse representation provides a new way to solve the above problems. The significance of sparse representation is that the signal can be represented in a given sparse dictionary as a linear combination of a small number of atoms [3-5]. The sparse representation method can effectively reduce the redundant components in the signal and extract mechanical weak fault features with finite data points. One problem of sparse representation is how to design an effective sparse representation basis. At present, there are some fixed basis, such as discrete cosine transform (DCT) basis and Gabor basis [3,4]. The basis can achieve good sparse effect for simple and stable signals, but show poor sparse representation effect for complex noisy signals. However, actual signals are diverse. Consequently, an appropriate and universally applicable sparse representation basis is particularly vital [6]. Recently, the sparse basis has gradually changed from fixed basis to learning basis. We update and construct a sparse representation basis which suitable for a specific signal by learning the feature information of the signal. Common learning algorithms of basis include the K-singular value decomposition (K-SVD) [7] and the method of optimal directions (MOD) [8]. During each iteration, the MOD method updates the whole basis atom, which requires large computation, but the K-SVD algorithm only updates one atom. When the signal to noise ratio (SNR) is low, other interference components, such as noise will produce a great influence on the signal, which makes the basis also learn those components, resulting in poor sparse representation.

The above dictionary learning methods all adopt the orthogonal match pursuit (OMP) algorithm [8] to update the sparse coefficients. This algorithm is realized by optimizing l_0 norm, which is a nonconvex penalty. OMP algorithm is a greedy algorithm, which is prone to over-fitting when the amount of data is large. In 2010, the online dictionary learning (ODL) is proposed to overcome those shortcomings [9]. The ODL utilizes the least absolute shrinkage and selection operator (*LASSO*) method to optimize l_1 to ensure the sparsity of the model, instead of l_0 [9], which avoids the problem that the sparse representation is easy to produce local optimal solution. The basis updates step adopting gradient descent method, which makes the algorithm faster and easier to find the deep features of the signal.

In order to improve the anti-noise ability of basis atoms, a new feature enhanced method based on the empirical wavelet transform (EWT) and sparse representation is

---

proposed in this paper. Gaussian filter is used to process the signal spectrum, reducing spectrum noise while retaining signal characteristics. Next, the empirical wavelet transform is utilized to divide the frequency spectrum into different frequency bands. The kurtosis value is used to select the optimal section, which is prepare for learn a basis. Gradient descent method is applied to train a basis. In the iterative process, the singular value of the basis matrix is processed by soft threshold to get a purer basis matrix. This idea comes from the singular value threshold (SVT) [10] and the updating method of ODL basis matrix. The LASSO with elastic net is applied to represent signal. Then the fault features are detected by envelope spectrum.

The following is the structure of this paper. The second part is the theoretical basis, the third part introduces the experimental process, the fourth part is the simulation experiment and experimental signal verification, and the conclusion remarks are drawn in the fifth part.

II. BASIC THEORY

According to the typical sparse representation framework, there are two problems to be solved:

A. Disadvantages of Traditional Sparse Representation:

1) Sparse representation

The sparse representation is proposed by Mallet, according to the theory, the signal can be represented by a complete basis [6]. The algorithm can be expressed as:

$$\arg\min_{X} \frac{1}{2}\|Y - DX\|_F^2$$
$$st\|X\|_0 \leq \omega \tag{1}$$

where, ω is the sparsity, $Y = \{y_i\}_{i=1}^n, Y \in R^{m \times n}$ is the signal matrix, $D \in R^{m \times m}$ is the basis, $X = \{x_i\}_{i=1}^n, X \in R^{m \times n}$ is the sparse coefficients matrix, $\|\ \|_0$ is l_0 norm, which is the sum of nonzero numbers of signal, $\|A\|_F$ is $\sqrt{tr(A^H A)}$ and $tr()$ is the sum of diagonal elements of a matrix. This algorithm trends to occur over-fitting, which may cause the fault feature of the signal not to be retained. It is critical aspect in the sparse algorithm to select a more matched sparse coding.

2) The construction of the basis matrix

In traditional dictionary learning, the update of the basis matrix can be viewed as achieving $\arg\min_{D}\|Y - DX\|_F^2$. The MOD dictionary learning updates D by the least squares [8]. However, rotating machinery works in a complex condition, which will lead a lot of redundant information. For the reason, the traditional method is more likely to learn the false atom, which with useless information, and lead the failure of the sparse representation. Thus, how to learn the sparse basis, which can best match the signal, also is the key of signal sparse representation.

B. The Method Proposed in This Paper:

1) Sparse representation:

In order to avoid the shortcomings of greedy algorithm, this paper chooses LASSO with elastic net to update sparse coefficients.

The LASSO algorithm can be expressed as the following equation [9]:

$$\arg\min_{X}\|Y - DX\|_F^2 + \lambda\|X\|_1 \tag{2}$$

where, $\|\ \|_1$ is sum of absolute values, and λ is the threshold, the LASSO can be regarded as a soft threshold processing [9].The LASSO with elastic net used in this study can be described by:

$$\arg\min_{X} \frac{1}{2}\|Y - DX\|_F^2 + \lambda_1\|X\|_1 + \lambda_2\frac{1}{2}\|X\|_F^2 \tag{3}$$

where, λ_1 and λ_2 are hyper-parameter, λ_1 controls the signal sparsity .This paper solves (3) by the least angle regression (LARS) [9].

2) Construction of the basis matrix:

EWT can be a powerful signal analysis method, which is proposed by Gills in 2013 [11]. It based on the Meyer wavelet to extract the signal information, EWT can be understood as a band pass filter group based on the frequency band, the fourier domain is defined in $[\ 0\ ,\pi\ ]$, as shown in Fig 1.
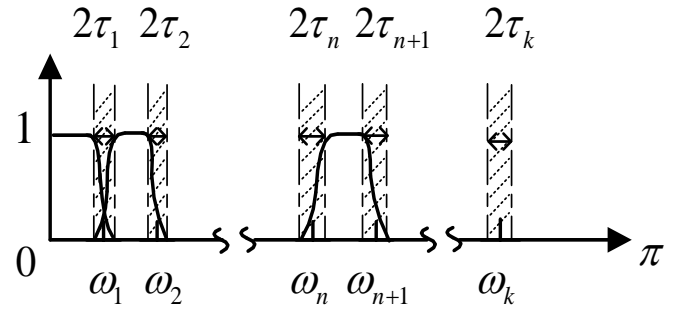
Fig. 1.Standardization of fourier spectrum

where, ω_n is the intermediate value of the adjacent maximum value, the $2\tau_n$ is transition phase.

In order to make the spectrum segmentation more accurate, this paper proposed a filter method based on gaussian blur, which is commonly applied in the field of image processing to make the image smoother by blurring the image boundary. It can be regarded as a convolution process. The Gaussian filter is as follows:

$$I_i = I * G_\sigma$$
$$G_\sigma = \frac{1}{\sigma\sqrt{2\pi}} e^{(x^2+y^2)/2\sigma^2} \tag{4}$$

where, I is the filtered signal, I_i is the filtered result, G_σ is the filter, σ is variance.

The spectrum after gaussian blur is divided into different N parts by selecting the local maximum value. The boundary of each part is finally used for the original signal frequency band [11].

The experience wavelet is defined a band pass filter group [11]. The expression of Meyer wavelet is as follows:

scaling function:

scale function:

$$\hat{\phi}_n(\omega) = \begin{cases} 1 & if \ |\omega| \le \omega_n - \tau_n \\ \cos[\frac{\pi}{2}\alpha(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n))] \\ & if \ \omega_n - \tau_n \le |\omega| \ge \omega_n + \tau_n \\ 0 & otherwise \end{cases} \quad (5)$$

wavelet function:

$$\hat{\psi}(\omega) = \begin{cases} 1 & if \ \omega_n + \tau_n \le |\omega| \ge \omega_{n+1} - \tau_{n+1} \\ \cos[\frac{\pi}{2}\alpha(\frac{1}{2\tau_{n+1}}(|\omega| - \omega_{n+1} + \tau_{n+1}))] \\ & if \ \omega_{n+1} - \tau_{n+1} \le |\omega| \ge \omega_{n+1} + \tau_{n+1} \\ \sin[\frac{\pi}{2}\alpha(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n))] \\ & if \ \omega_n - \tau_n \le |\omega| \ge \omega_n + \tau_n \\ 0 & otherwise \end{cases} \quad (6)$$

where, $\alpha(x)$ can be expressed as the following function:

$$\alpha(x) = \begin{cases} 0 & if \ x < 0 \\ \alpha(x) + \alpha(1-x) = 1 & \forall x \in [0,1] \\ 1 & if \ x \ge 1 \end{cases} \quad (7)$$

where, $\alpha(x)$ is a smooth function [11], which selects the most common functions as:

$$\alpha(x) = x^4(35 - 84x + 70x^2 - 20x^3) \quad (8)$$

the low frequency information is obtained by extracting the convolution part of the scale function and signal, as well as the wavelet function and signal convolution to obtain the high frequency part as shown in (9) :

$$W_f^\varepsilon(0,t) = \left\langle \hat{f}, \hat{\varphi} \right\rangle \ , \ W_f^\varepsilon(n,t) = \left\langle \hat{f}, \hat{\psi} \right\rangle \quad (9)$$

where, $W_f^\varepsilon(0,t)$ is the low frequency part, $W_f^\varepsilon(n,t)$ is the high frequency part, $\hat{f}$ is the spectrum of signal. By band pass filtering in the frequency domain of the signal, we can obtain the components which contain the feature of vibration signal.

The all parts contain all the information of the signal. Using different parts for learning, the fault information hidden in different frequency bands can be obtained. We can utilize characteristics of bearing signal to select the optimal part, such as kurtosis [9]. At the same time, the energy of fault signal in fourier spectrum is relatively concentrated. We can learn an optimal basis by using the optimal part.

The basis learning is based on the following:

$$\arg\min_D \|Y - DX\|_F^2 \quad (10)$$

the sparsity coefficients X can be solved by *LASSO* with elastic net, this equation can be rewritten as:

$$\arg\min_D \|Y - DX\|_F^2$$
$$= \left\| (Y - \sum_{j \ne k}^n d_j x_j^T) - d_k x_k^T \right\|_F^2 = \left\| E_k - d_k x_k^T \right\|_F^2 \quad (11)$$

where, E_k is residual matrix. In order to ensure the sparsity, we only select the x_k^T nonzero columns in E_k, which can also reduce useless information. The gradient descent method is used to decompose the E_k. The gradient descent can be written as follows:

$$D_n = D_{n-1} - \rho \frac{\partial E_k}{\partial D} \quad (12)$$

$\left\| E_k - d_k x_k^T \right\|_F^2$ can be rewritten as $E_k \approx d_k x_k^T$,where, $E_k \subset R^{m*n}, d_k \subset R^{m*1}, x_k^T \subset R^{1*n}$ .The loss function is written as:

$$\left\| E_k - d_k x_k^T \right\|_F^2 = tr((E_k - d_k x_k^T)^T(E_k - d_k x_k^T)) \quad (13)$$

the partial derivative of (13) is shown in (14):

$$\frac{\partial tr((E_k - d_k x_k^T)^T(E_k - d_k x_k^T))}{\partial d_k} = 2(E_k - d_k x_k^T)x_k \quad (14)$$

update d_k by gradient descent:

$$d_k^n = d_k^{n-1} - (2\rho(E_k - d_k^{n-1} x_k^T)x_k) \quad (15)$$

the learning rate ρ is set as:

$$\rho = d_k^{n-1}./2d_k^{n-1} x_k^T x_k \quad (16)$$

The learning rate is a positive number to ensure the convergence of the function.

The above methods are applied to update the basis and get sparse coefficients. The final iteration function can be written as:

$$d_k^n = d_k^{n-1}. \times E_k x_k / d_k^{n-1} x_k^T x_k$$
$$x_k^{T\ n} = x_k^{T\ n-1}. \times d_k^T E_k./ d_k^T d_k x_k^{T\ n-1} \quad (17)$$

Finally, we deal with the singular value of the basis matrix by soft threshold in the iterative process and retain the part with the maximum energy. The algorithm is equal to a constraint on the rank of the matrix as shown in:

$$D_n = \arg\min \frac{1}{2}\|D_n - D\|_F^2 + rank(D_n) \quad (18)$$

where, rank penalty is nonconvex penalty, so we rewrite this function as:

$$D_n = \arg\min \frac{1}{2}\|D_n - D\|_F^2 + \zeta\|D_n\|_* \quad (19)$$

$\| \ \|_*$ is the nuclear norm, which is the sum of all singular values, and the nuclear norm is a convex constraint. ζ is a

parameter which control the rank. This algorithm selects large singular values to reduce atomic noise. The proximal operator is as follows [10]:

$$[U, \Lambda, V] = \text{svd}(D)$$
$$D_n = V * \max(0, \Lambda - \zeta I) * U^T \qquad (20)$$

where, U is the right hand matrix of SVD, V is the left hand matrix of SVD, Λ is the singular value. Then a pure basis can be obtained from the above algorithm.

The basis matrix is updated as follows:

Basis matrix update algorithm
Input: initial basis D ,signal Y ,n=1, iter =20
while the number of iterations is not reached

$1. X = \arg\min \|Y - DX\|_F^2 / 2 + \lambda_1 \|X\|_1 + \lambda_2 \|X\|_F^2 / 2$

$2. \|E_k - d_k x_k^T\|_F^2 \rightarrow d_k^n = d_k^{n-1}. \times E_k x_k ./ d_k^{n-1} x_k^T x_k \rightarrow D$

$\qquad x_k^{T\,n} = x_k^{T\,n-1}. \times d_k^T E_k ./ d_k^T d_k x_k^{T\,n-1} \rightarrow X$

$3. [U, \Lambda, V] = \text{svd}(D)$

$4. D_n = V * \max(0, \Lambda - \zeta I) * U^T$

$5. normalize(D)$

$6. n = n + 1$

end
Output:D

III. OPERATION FLOW OF THE WEAK FAULT EXTRACTION METHOD

In industrial production, if the bearing model is known, the fault frequency can be calculated by:

$$f_r = \frac{Z}{2}\left(1 - \frac{d}{D}\cos\theta\right)f_0 \qquad (21)$$

where Z is the number of roller elements, f_r is the fault frequency, d is roller diameter, D is pitch diameter and θ is contact angle, f_0 is the rotational frequency of shaft.

In order to verify the superiority and effectiveness of this method, the simulation signals and experimental signals are utilized. The process of this paper is as shown in Fig. 2
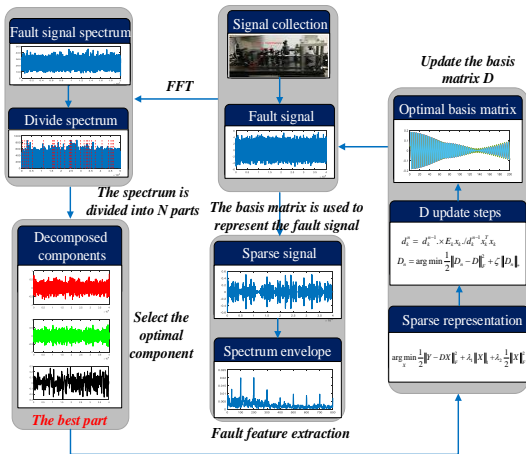
Fig. 2. The method flowchart of this method.

The following is the process of this method:

(1) The frequency band is segmented by Gaussian filter and EWT. The Gaussian filer can avoid the abnormal points of spectrum, the EWT is employed to process the signal spectrum to get different IMF.

(2) The information of different frequency band is extracted and the optimum part is selected by kurtosis.

(3) The selected part is utilized to learn the sparse basis, while the sparse signal can be obtained from the original signal and sparse basis.

(4) Finally, we extract features from the envelope spectrum of sparse signals.

IV. APPLICATION CASES

In the contrast test, the basis was learned by K-SVD, and then we use this basis to sparse represent the fault signals. The LASSO with elastic net is used to sparsely represent the fault signal. Gaussian filter variance is 1.5, the length of filter is 500.

$$y(t) = Ae^{-2\pi f_n \xi t} \times \sin\left[2\pi f_n t \sqrt{1-\xi^2}\right] + v(t) \qquad (22)$$

where, A is the amplitude and ξ is the damping coefficient, f_n is the natural frequency, $v(t)$ is the noise. The parameters of simulation signal are as follows: $f_n = 25000$ (Hz), $\xi = 0.06$, $A = 3$. The simulation time domain waveform and envelope spectrum are shown in Fig.3, where the fault frequency of the simulation signal is 100 (Hz).
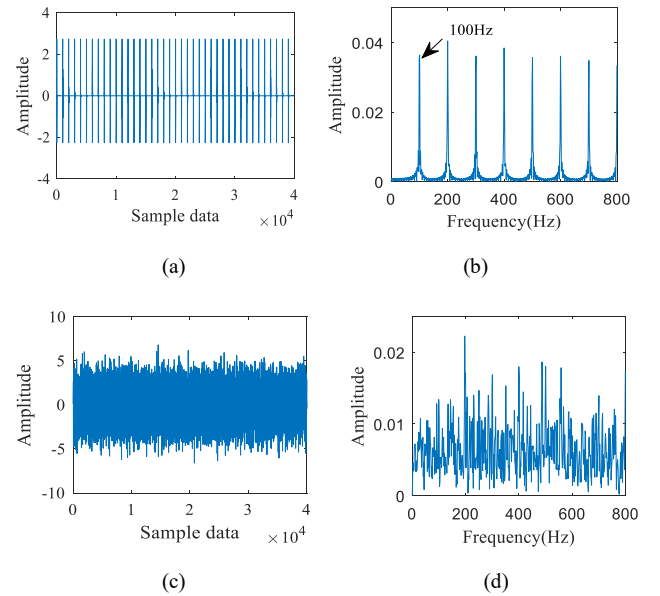
Fig.3. The simulation signal and the signal envelope spectrum. (a) The simulation time-domain waveform. (b) The envelope spectrum of simulation signal. (c) Time domain diagram of simulation signal with -3db noise added. (d) The envelope spectrum of noise signal.

The frequency spectrum is divided into 30, and the divided frequency bands are shown in Fig.4 (a). After that, the EWT is used to obtain the intrinsic mode function (IMF). Fig.4 (b) is the kurtosis of each IMF. (c) is the time domain of optimal IMF, (d) is the optimal envelope spectrum of IMF.
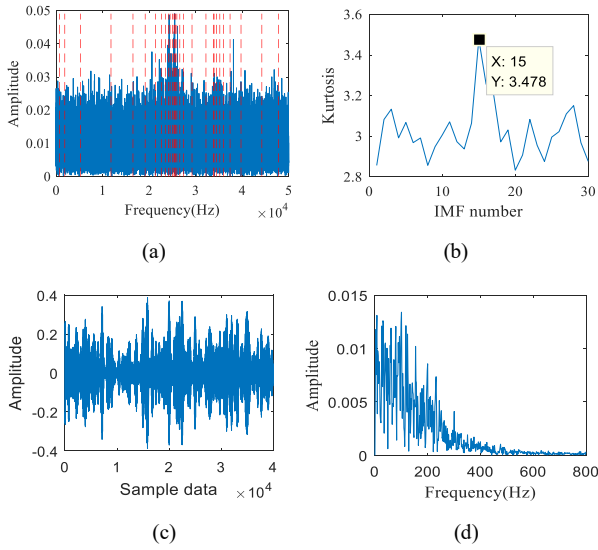
Fig.4. Spectrum division and optimal mode selection. (a) The divided frequency bands. (b) The kurtosis of each IMF. (c) The time domain of optimal IMF. (d) The optimal IMF envelope spectrum.

Next, we use the part with the maximum kurtosis value as the new signal and train a basis using the method in this paper. Fig.5 (a) shows the basis learned by K-SVD dictionary learning, Fig.5 (b) is the basis learned by this method. Among them, the basis learned by K-SVD dictionary has a large noise, and the basis learned by the method in this paper has an obvious impact component, where parameters are as follow: $\lambda_1 = 0.1, \lambda_2 = 0.3, \zeta = 0.08$.
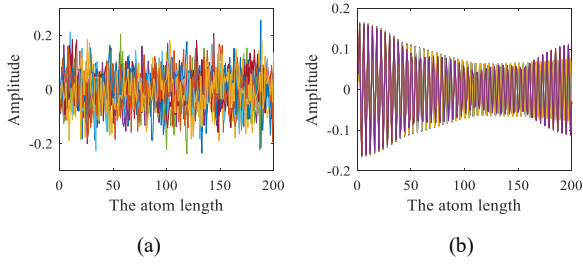
Fig.5. Basis matrix. (a) The basis learned by K-SVD dictionary learning. (b) The basis learned by this method.
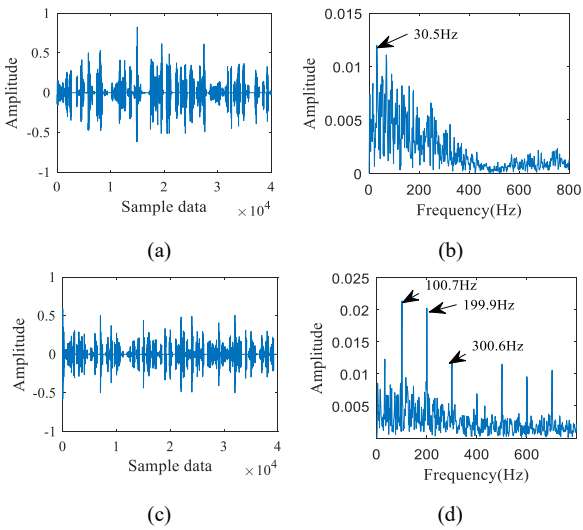
Fig.6. Sparse signal and sparse signal envelope spectrum. (a) Sparse signal obtained by K-SVD. (b) The envelope spectrum of sparse signal based on K-SVD. (c) Sparse signal obtained by this method. (d) The envelope spectrum of sparse signal based on this method.

Next, the *LASSO* with elastic net is used to sparsely represent the original signal through the obtained basis. To ensure the reliability of experimental comparison, we adopted the sparsity of sparse signals as the discriminant standard.

The rationality of the experiment can be guaranteed by the same sparsity. The basis learned by K-SVD is used for sparse representation of the original signal as shown in the Fig.6 (a). The fault feature can be extracted from envelope spectrum Fig.6 (b). The basis learned by this method is utilized for sparse representation effect as shown in Fig.6 (c), the envelope spectrum shown in (d).

To prove the feasibility of the proposed method in practical engineering application, bearing with outer race fault is employed for feature extraction. The indentation of the width of 0.7 mm and the depth of 0.25 mm in the outer-race of the bearing. The shaft speed was 1300 (rpm) and a larger sampling frequency is selected to obtain more comprehensive mechanical state information. The sampling frequency of the bearing experiment is 100 (kHz). The outer-race fault frequency is 86.32 (Hz). The experimental system of this paper is shown in Fig. 7. The time domain waveform and envelope spectrum of the signal are shown in Fig.8.
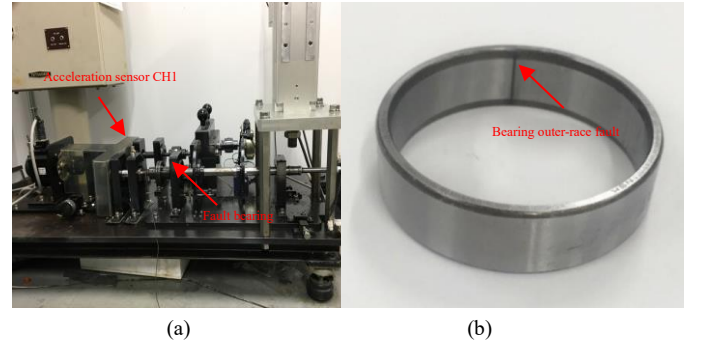
Fig.7. Experimental system. (a) Experimental table. (b) Bearing outer-race fault.
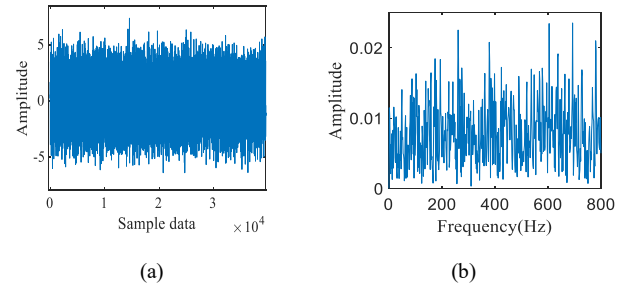
Fig.8. Outer-race bearing signal and envelope spectrum. (a) The out race time-domain waveform. (b) The envelope spectrum.
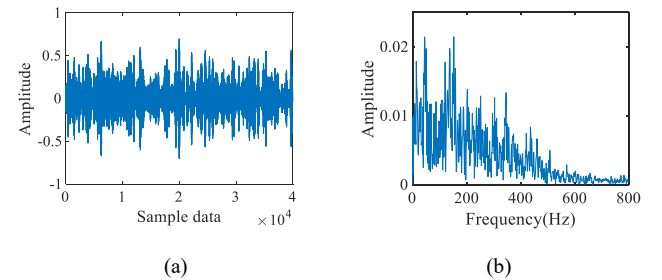
Fig.9. (a) The time domain of optimal IMF. (b) The optimal IMF envelope spectrum.

528
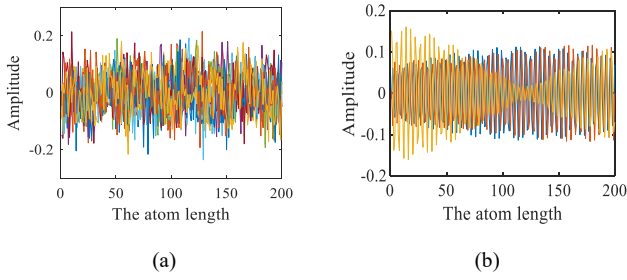
Fig.10. Basis matrix. (a) The basis learned by K-SVD dictionary learning. (b) The basis learned by this method.
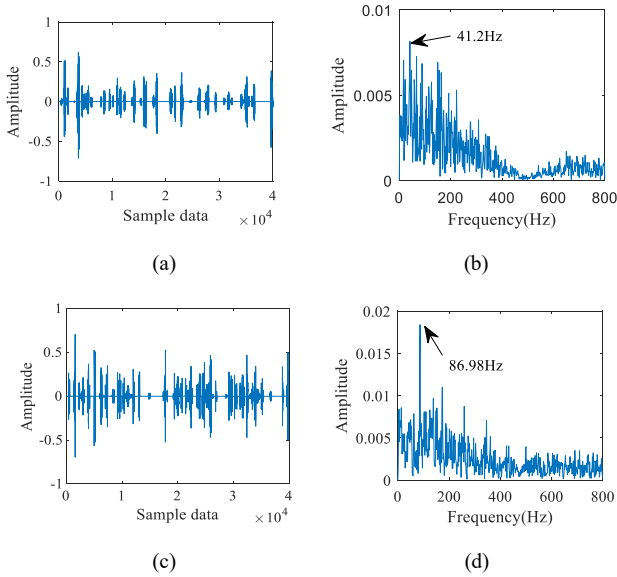
Fig.11. Sparse signal and sparse signal envelope spectrum. (a) Sparse signal obtained based on this K-SVD. (b) The envelope spectrum of sparse signal based on this K-SVD. (c) Sparse signal obtained based on this method. (d) The envelope spectrum of sparse signal based on this method.

The part with the maximum kurtosis, is selected to learn a basis, the chosen IMF is shown in Fig.9 (a), The envelope spectrum of this signal is shown in Fig.9 (b).

The basis learned by K-SVD dictionary learning and the basis learned by this method are shown in Fig.10. The obtained basis is used for sparse representation of the original signal, sparse signal and spectral envelope are shown in Fig. 11. It is proved from the envelope spectrum that this method has better feature extraction ability than the traditional method.

V. CONCLUSION

In order to accurately extract the weak fault feature of rolling bearing. This paper proposes a novel weak fault extraction method based on sparse representation and EWT. gaussian filter is used to regularize the spectrum so that the division of fault signal spectrum division can be more accurate. The optimal part is selected through kurtosis and it is used to learned the basis matrix. Then, the original signal is sparsely represented by the basis matrix. Finally, the fault feature is extracted by envelope spectrum. Experimental results show that the method proposed in this paper can retain the feature of the original signal better than the traditional method.

REFERENCES

[1] Y. Hao, L. Song, M. Wang, L. Cui, and H. Wang, "Underdetermined source separation of bearing faults based on optimized intrinsic characteristic scale decomposition and local non-negative matrix factorization," IEEE Access, vol. 7, pp. 11427-11435, 2019.

[2] L. Cui, B. Lei, J. Ma, and Z. Jin, "Quantitative trend fault diagnosis of a rolling bearing based on Sparsogram and Lempel-Ziv". Measurement, vol. 128, pp. 410-418, 2018.

[3] L. Wang, G. Cai, and W. You "Transients Extraction Based on Averaged Random Orthogonal Matching Pursuit Algorithm for Machinery Fault Diagnosis". IEEE Transactions on Instrumentation & Measurement, vol. 66, pp3237-3248,2017

[4] H. Wang, Y. Ke, and L. Song, "A Sparsity-Promoted Decomposition for Compressed Fault Diagnosis of Roller Bearings". Sensors, vol.16, 2016

[5] L. Song, H. Wang, and P. Chen, "Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery," IEEE Transactions on Instrumentation & Measurement, vol. 67, pp. 1887-1899, 2018.

[6] B. Ren, Y. Hao, H. Wang, L. Song, G, Tang, and H, Yuan, "A sparsity-promoted method based on Majorization-Minimization for weak fault feature enhancement". Sensors, vol. 18, 1003, 2018.

[7] H. Yang, H. Lin, and K. Ding. "Sliding window denoising K-Singular Value Decomposition and its application on rolling bearing impact fault diagnosis". Journal of Sound & Vibration vol. 421, pp. 205-219, 2018.

[8] C. Liu, X. Wu, and J. Mao "Sparse representation of rolling bearing vibration signal based on improved MOD dictionary learning". Computer Engineering and Applications, vol 52: pp 257-260, 2016.

[9] H. Wang, P. Wang, L. Song, B. Ren, and L. Cui "A Novel Feature Enhancement Method Based on Improved Constraint Model of Online Dictionary Learning". IEEE Access, vol. 7, pp.17599 – 17607,2019.

[10] J .Cai, E. Candes, and Z Shen, "A Singular Value Thresholding Algorithm for Matrix Completion". SIAM Journal on Optimization .pp.1956-1982, 2010.

[11] J. Gilles, "Empirical Wavelet Transform". IEEE transaction on signal processing, vol.61, pp.3999-4010, 2013.

A Weak Fault Diagnosis Method Based on Sparsity Overlapping Group Lasso for Rolling Bearing

Li Qiu
School of mechanical and electrical engineering
Beijing University of Chemical Technology
Beijing, China
1788882190@163.com

Xiaoming Wang
School of mechanical and electrical engineering
Beijing University of Chemical Technology
Beijing, China
18794874195@163.com

Zhengcai Guo
School of mechanical and electrical engineering
Beijing University of Chemical Technology
Beijing, China
buctgzc@163.com

Guoan Yang*
School of mechanical and electrical engineering
Beijing University of Chemical Technology
Beijing, China
yangga@mail.buct.edu.cn

Abstract—**The rolling bearing is a vital part of modern equipment. It has great significance to effectively extract fault periodic pulses from bearing signals. Due to the noise mixed in bearing signal, the extraction of weak fault feature is facing difficult. For this, a new method for diagnosis of rolling bearing is developed in this paper, based on weighted sparsity norm and overlapping group sparse/shrinkage (OGS), which called sparsity overlapping group lasso (SOGL).The signal obtained from working rolling bearing is prone to be interfered with noise and occur coupling in time domain, so the algorithm processes signal in the Fourier domain, and combines with iterative shrinkage threshold (ISTA) and majorization–minimization (MM) to solve the model proposed in this study. Because signal mixed with noise will reduce the processing accuracy, based on the variational mode decomposition (VMD), this paper derives a simple and fixed center frequency filtering algorithm to reduce the useless components in the signal and improve the effect of SOGL. The SOGL is applied to simulated signals and measured signals. From the experimental results, the weak fault periodic pulse obtained by SOGL is obviously prominent, and this method has better performance in feature extraction comparing with the method of fast Fourier transform and OGS (FFT-OGS).**

Keywords—overlapping group Lasso, Sparse representation, fast Fourier transform

I. INTRODUCTION

The condition monitoring and diagnostic technology have been widely applied in the field of machinery equipment[1]. Vibration signals from working mechanical equipment contain a lot of useful information, so vibration analysis is a common signal processing technique. As a vital part of mechanical equipment, fault diagnosis is pivotal for the bearing. While it is running in a bad environment, the local defects (such as wear, pitting and so on) are prone to occur [2]. The machinery becomes higher complexity, thus the diagnostic methods based on vibration have been facing more and more challenge [3]. Suffering from various factors, the vibration signal of bearing is nonlinear, which will make the fault feature be submerged [4]. For bearing with weak fault, common signal processing methods cannot directly extract features from vibration signals.

Sparse representation has been widely applied in signal denoising [5], which originates from atomic decomposition [6] and can effectively analyze the potential characteristics of the signal. In 1993, matching pursuit (MP) [7] was presented to deal with sparse atomic decomposition, which finds the atom with the largest residual error compared to the original signal and calculates its corresponding sparse coefficient. After that, many algorithms based on MP were successively developed. The orthogonal matching pursuit (OMP) [8-9] was studied in 1993–2007, which ensures that the residual after each selection is orthogonal to the atom, thus speeding up the iteration speed. The model of the least absolute shrinkage and selection operator (Lasso) is proposed in 1996, which transforms the problem based on l_0 norm into a regularization optimization solution based on l_1 norm [10]. They also proved that these two sparse norms were equivalent under certain circumstances, and the Lasso has stronger feature selection ability comparing with traditional greedy algorithms such as MP and OMP. In 2006, Huan [11] proposed efficient algorithms for the extensions of Lasso called group Lasso, which adopts group l_2 norm as constraint and forces sparsity between groups. To date, group sparsity has been a hot research area in signal processing, machine learning. An overlapping group shrinkage (OGS) algorithm to process speech signal is studied in 2014[12]. And zhao [13] proposed an method called period group lasso (PGL), which sets the group periodic. The method can be applied to extract bearing composite fault feature.

The above methods all process signals based on time domain. Due to the serious coupling between the actual signal impact and noise, the traditional OGL and PGL are difficult to effectively separate the shock information. For this kind of signal, it is necessary to decouple the signal in advance, so as to realize the separation of shock and noise.

In this paper, an improved denoising model based on weighted l_1 norm and OGS is studied for better extracting weak fault of bearing, which called sparsity overlapping group lasso (SOGL). The method combines with the fast iterative shrinkage threshold (ISTA) and soft threshold operations(STO) for optimization of sparsity model. The SOGL processes signal in the Fourier domain to get denoised one, next the inverse Fourier transform will be used for it. The real part of signal is screened out and the fault type is judged by envelope spectrum. To improve the processing accuracy of this method, based on the variational mode decomposition (VMD), we derive a fixed center frequency

---

filtering algorithm as pre-processing, to reduce the useless components in the signal.

The remainder of this paper is organized as follows: Section 2 briefly introduces the relevant theoretical basis and describes our method in details. In Section 3, we show the flow of this algorithm and apply the method for simulation signal. Some real experimental data from bearing are used to verify the effectiveness of the proposed method in Section 4. Finally, conclusions are given in Section 5.

II. THEORETICAL BASIS

A. Sparse representation and the OGS

The sparse representation is described as:

$$\arg\min_X \|Y - DX\| + \phi(X) \tag{1}$$

where, the $\phi(X)$ is the sparse penalty term, $Y \in R^{n\times 1}$ denotes the original signal, $X \in R^{n\times 1}$ represents the output signal. And the $\phi(X)$ usually is $\|\ \|_p^p$ sparse norm or $sum\|\ \|_2$ sparse group penalty.

The OGS can effectively detect waveform energy and filter out noise. The nonconvex sparsity OGS is described in follows:

$$\arg\min_X \frac{1}{2}\|W - X\|_2^2 + \lambda\varphi(sum\|x_{i,k}\|_2) \tag{2}$$

where, $\varphi(x)$ represents the nonconvex sparse promoter, $x_{i,k} = [x_i, \dots x_{i+k-1}]$, λ is sparse balance parameter, problem(2) be solved by majorization–minimization (MM) method, the iteration of MM can be described as:

$$x_{n+1} = \arg\min_x F(x, x_n) \tag{3}$$

where, $F()$ is a majorizer of $\varphi()$. The (3) satisfies $F(x, x_n) > \varphi(x), F(x_n, x_n) = \varphi(x)$. The Taylor expansion of nonconvex constraint $\varphi(x)$ as follows:

$$\varphi(x) = \varphi(v) + \nabla\varphi(v)(x - v) + \frac{\nabla^2\varphi(v)}{2}\|x - v\|_2^2 \tag{4}$$

Where, $\frac{\nabla^2\varphi(v)}{2}\|x - v\|_2^2$ denotes the higher order infinitesimal term, v represents the value of the last iteration, $\varphi(x) \le \varphi(v) + \nabla\varphi(v)(x - v)$, because $x \le \frac{1}{2v}x^2 + \frac{v}{2}$, the

$\varphi(x) = \varphi(v) + \nabla\varphi(v)(\frac{1}{2v}x^2 - \frac{v}{2})$.

The (2) can be redefined as:

$$\arg\min_x 1/2\|y - x\|_2^2 + \lambda(sum(\frac{\varphi'(\|v_{i,k}\|_2)}{\|v_{i,k}\|_2} \times \|x_{i,k}\|_2^2))) + C \tag{5}$$

The fixed point iteration form of this operator is as follows:

$$x_n = \frac{y}{1 + \lambda\dfrac{\varphi'(\|x_{n-1,i,k}\|_2)}{\|x_{n-1,i,k}\|_2}} \tag{6}$$

B. VMD

The model of VMD can be described as follows,

$$\begin{cases} \min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\|\partial_t\left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 \right\} \\ s.t. \quad \sum_k u_k = f \end{cases} \tag{7}$$

Where, u_k and ω_k denote the modes and their center frequencies.

The (7) be changed into unconstrained one by applying quadratic penalty parameter α and Lagrange multiplier $\lambda(t)$, thus this problem can be redefined as:

$$L(\{u_k\},\{\omega_k\},\lambda(t)) = \alpha\sum_k \left\|\partial_t\left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 \tag{8}$$

$$+ \left\|f(t) - \sum_k u_k(t)\right\|_2^2 + \left\langle\lambda(t), f(t) - \sum_k u_k(t)\right\rangle$$

The k-th component u_k and its center frequency ω_k are updated by (9)

$$\begin{cases} \hat{u}_k^{n+1}(\omega) = \dfrac{\hat{f}(\omega) - \sum_{i\neq k}\hat{u}_i(\omega) + \dfrac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \\ \omega_k^{n+1} = \dfrac{\int_0^\infty \omega|\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \end{cases} \tag{9}$$

Where, $\hat{u}_k(\omega)$, $\hat{\lambda}(\omega)$ and $\hat{f}(\omega)$ denote the Fourier transform of mode u_k, Lagrange multiplier $\lambda(t)$ and the raw data $f(t)$.

C. The method proposed in this paper

1) The fixed center frequency filtering algorithm

In this paper, the input parameter K which determines the number of decomposition modes in VMD is fixed to 1, which can retain the full frequency band of bearing signal and remove some redundant components. The quadratic penalty parameter α is set as 500. We set the update value of VMD center frequency to be fixed, which speeds up the iteration. The VMD becomes a Wiener filter with a fixed center frequency, which can effectively filter out some useless frequency components.

The modal of fixed center frequency filtering algorithm as follows:

$$L(f'(t),\lambda(t)) = \alpha\sum_k \left\|\partial_t\left[\left(\delta(t) + \frac{j}{\pi t}\right) * f'(t)\right]e^{-j\omega_k t}\right\|_2^2 \tag{10}$$

$$+ \|f(t) - f'(t)\|_2^2 + \langle\lambda(t), f(t) - f'(t)\rangle$$

2) The SOGL based on Fourier domain

$$\arg\min_x 1/2\|y - ifft(x)\|_2^2 + \lambda(sum\|x_{i,k}\|_2^2) + \|D \odot x\|_1 \tag{11}$$

The $ifft()$ can be seen as A, then $fft()$ is regarded as A^H. The (11) is redefined as:

$$\arg\min_x 1/2\|y - A(x)\|_2^2 + \lambda(sum\|x_{i,k}\|_2^2) + \|D \odot x\|_1 \quad (12)$$

The problem (12) can be solved by iterative shrinkage threshold (ISTA). The Taylor expansion of arbitrary function $f(x)$ as follows:

$$f(x) = f(v) + \nabla f(v)(x - v) + \frac{L}{2}\|x - v\|_2^2 + g(x)$$
$$\rightarrow \frac{1}{2}\left\|x - v + \frac{\nabla f(v)}{L}\right\|_2^2 + \frac{1}{L}g(x) \quad (13)$$

Where, $g(x)$ denotes the constraint, L is Lipschitz constant and usually less than $\|A^H A\|_2$. In this paper, we carry out standard Fourier transform, and the value of Lipschitz constant is smaller than 1.

The optimization process of (13) is as follows:

$$z = v - \frac{\nabla f(v)}{L}$$
$$\arg\min_x \frac{1}{2}\|x - z\|_2^2 + \frac{1}{L}g(x) \quad (14)$$

Similarly, there is a optimized model of (12) as:

$$\arg\min_x \frac{1}{2}\|x - z\|_2^2 + \lambda(sum(\frac{\varphi'(\|v_{i,k}\|_2)}{\|v_{i,k}\|_2} \times \|x_{i,k}\|_2^2))) + \lambda_1\|D \odot x\|_1 \quad (15)$$

Since the above formula is a convex optimization, the following formula can be obtained by derivation:

$$(1 + \lambda\frac{\varphi'(\|x_{n-1,i,k}\|_2)}{\|x_{n-1,i,k}\|_2})\hat{X} = \hat{W} - \lambda_1 D(sign(\hat{X}))$$

$$\rightarrow \hat{X} = \hat{W}./(1 + \lambda\frac{\varphi'(\|\hat{X}_{n-1,i,k}\|_2)}{\|\hat{X}_{,i,k}\|_2}) \quad (16)$$

$$-\lambda_2/(|\hat{X}_{n-1}|(1 + \lambda\frac{\varphi'(\|\hat{X}_{n-1,i,k}\|_2)}{\|\hat{X}_{n-1,i,k}\|_2})sign(\hat{X})$$

Where, $sign()$ is signum function, as for complex number, the signum function of it can be described as

$$sign(a + bi) = \frac{a}{\sqrt{a^2 + b^2}} + \frac{bi}{\sqrt{a^2 + b^2}}, D = 1/(|\hat{X}|) \quad (17)$$

$$\hat{X}_n = prox\_l_1(\hat{W}./(1 + \lambda X_N, \lambda_2 *1/(|\hat{X}_{n-1}|(1 + \lambda X_N)) \quad (18)$$

Where, $prox\_l_1$ is the soft thresholding operation,

$$X_N = \frac{\varphi'(\|\hat{X}_{n-1,i,k}\|_2)}{\|\hat{X}_{n-1,i,k}\|_2}.$$

III. THE FLOW OF SOGL

Rotating machinery usually work in a noisy environment, this may cause some early bearing failure information to be submerged in noise. In this paper, an improved denoising model based on OGS is proposed, which called SOGL.

A. The SOGL mthod

Fig.1 is the process of the SOGL method, the specific steps are listed as follows:

(1) Get the original bearing signal. Observe the basic information of the signal from the time domain waveform and spectrum.

(2) Employ the fixed center frequency filtering algorithm based on VMD to reduce the useless components in the signal.

(3) Processes the denoised signal in Fourier domain by the SOGL. The inverse Fourier transform is applied to signal obtained.

(4) Extract the real part of the signal and the judged fault type through envelope spectrum.
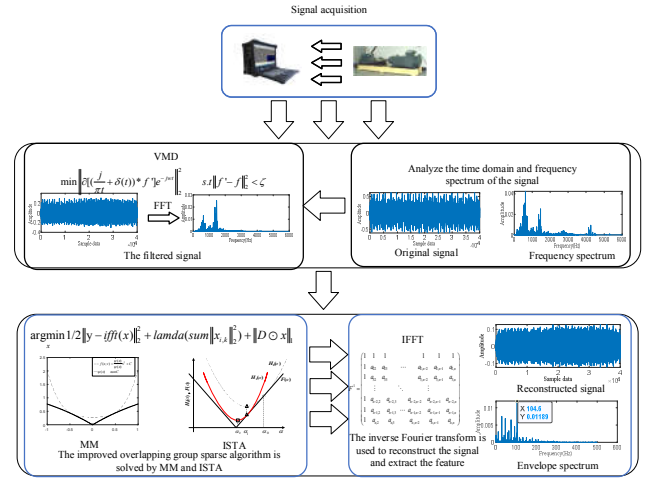
Fig. 1. The flow diagram of the fault feature extraction method based on SOGL

B. Simulation verification

Taking the characteristics of real signal into consideration, we construct the bearing fault simulation signal which with gauss white noise.

$$y = y_0 e^{-2\pi f_n \zeta t}\sin\left(2\pi f_n\sqrt{1 - \zeta^2}t\right) + n(t) \quad (19)$$

Where, y_0 and f_n respectively denote the amplitude of the signal and the natural frequency, ζ represents the damping coefficient.

And $n(t)$ is noise. In here, the parameters are set as $y_0 = 3$, $f_n = 15000$ (Hz), $\zeta = 0.1$.

As for the fault characteristic frequency, it is 100Hz. While the sample frequency is set as 10000 Hz, 40000 is the number of sampling points.

From the Fig.2 ,we can clearly see the time domain waveform and envelope spectrum of the above simulation signal.
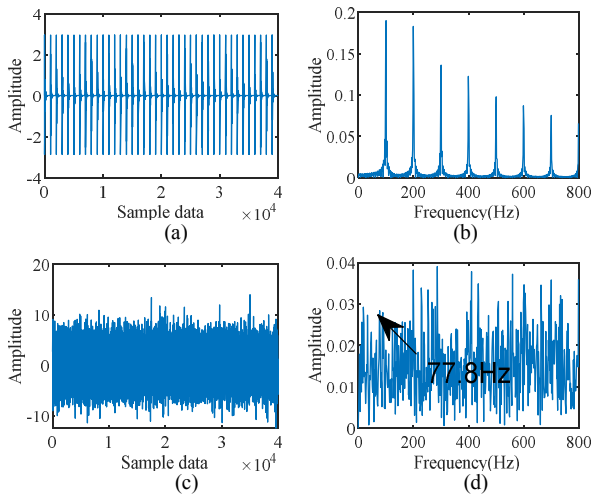
Fig. 2. The simulation signal and the signal envelope spectrum. (a) The simulation time-domain waveform. (b) The envelope spectrum of (a). (c) Time domain diagram of simulation signal with Gaussian noise added. (d) The envelope spectrum of noise signal.

The extraction of characteristic frequency is interfered by some useless frequency components.

The simulation signal is processed as follows: Firstly, the main parameters of filter like K and α are fixed as 1 and 500. Next transform the signal into frequency domain and the SOGL is utilized for further processing.

Then the inverse Fourier transform is used for transform signal into time domain. Extract the real part of the signal and the judged fault type through envelope spectrum.

Fig. 3. The reconstruct signal and its spectrum envelope spectrum. (a) Signal recovered by the method proposed in this paper. (b) The envelope spectrum of (a).

The result is as Fig.3. The failure frequency is obvious. The SOGL which adopts weighted l_1 norm and ISTA can reconstruct signal well and highlight the fault characteristic frequency.

IV. EXPERIMENTAL APPLICATION

Some experimental data from the bearing test bench is used to verify the validity and superiority of the SOGL algorithm. The rolling element bearing test bench is shown in Fig.4.
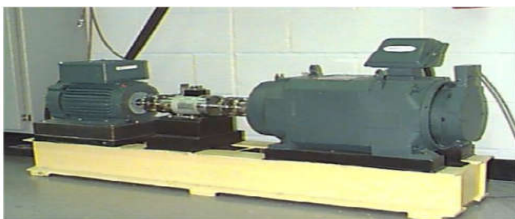
Fig. 4. Bearing test bench

Generally, if the bearing model and operating parameters are known, the failure frequency can be calculated by:

$$f_{out} = \frac{Z}{2}\left(1 - \frac{d}{D}\cos\theta\right)f_s$$
$$f_{in} = \frac{Z}{2}\left(1 + \frac{d}{D}\cos\theta\right)f_s \tag{20}$$

Where, the fault frequency of outer race is represented by f_{out}, the fault frequency of inner race is f_{in}, the number of roller elements is denoted by Z, d is the diameter of roller, D is pitch diameter, θ denotes the contact angle and f_s is the rotor mechanical frequency.

In this paper, two groups of bearing open data sets from Case Western Reserve University are selected. One group of data belong to the outer race fault of bearing, the other one is the inner race fault. The bearing model is 6205-2RS JEM SKF, which is a deep groove ball bearing, and the damage is point fault. The specific parameters of the bearing are shown in Table.

TABLE I. THE SPECIFIC PARAMETERS OF THE 6205-2RS JEM SKF

Type of fault	Z	d	D	f_s
Outer race fault	9	7.94mm	39.04mm	29.17Hz
Inner race fault	9	7.94mm	39.04mm	29.53Hz

The fault characteristic frequency of outer race fault bearing and inner race fault bearing can be obtained by (20), which is respectively 104.6Hz, 159.9Hz.

In addition, one group of comparative experiment was set up, which utilizes the method of fast Fourier transform and OGS (FFT-OGS). The OGS model in the FFT-OGS method is described in (5).

A. Case Study for defective bearing (Outer race fault)

From the envelope spectrum of the outer race signal in the Fig.5, we can see the rotor mechanical frequency is very clearly, and the useful frequency is submerged by other frequencies.

For this signal, we deal with it in the fellow steps: Firstly, the outer race signal is processed by the fixed center frequency filtering algorithm. The obtained signal is used as the input of the next stage. The SOGL and the FFT-OGS are utilized for dealing with the signal in the frequency domain, which can reduce the noise mixed in the signal. The reconstruct signal is transformed into time domain. Finally, select the real part of the signal and diagnose fault type through envelope spectrum.
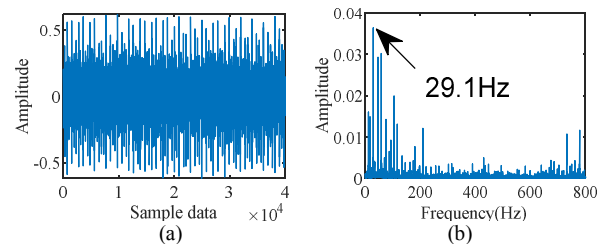
Fig. 5. The outer ring fault signal and its spectrum envelope spectrum. (a) The time-domain waveform of original signal. (b) The envelope spectrum of this signal.
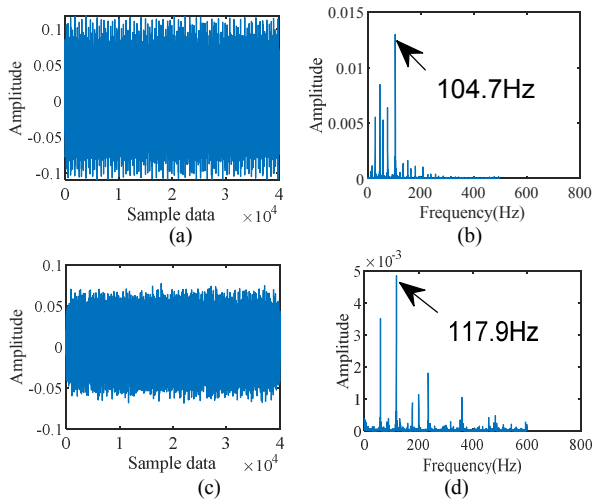
Fig. 6. The reconstruct signal and its spectrum envelope spectrum. (a) Reconstruct signal obtained from the method proposed in this paper. (b) The envelope spectrum of this reconstructed signal. (c)The signal obtained from FFT-OGL method. (d) The envelope spectrum of (c).

Fig.6 is the signal obtained from the method proposed in this paper and FFT-OGS algorithm. Comparing the results of the two methods, we found that the SOGL can well highlight the characteristic frequency and reduce the interference of other frequencies. However, the characteristic frequency in the signal which processed by the FFT-OGL method cannot be extracted well. The vital frequency is still submerged.

B. Case Study for defective bearing (Inner race fault)

Firstly, we observe the basic information of the signal. From the Fig.7, the failure frequency of 159.9Hz is not obvious at all, which can be regarded as a weak fault of rolling bearing.
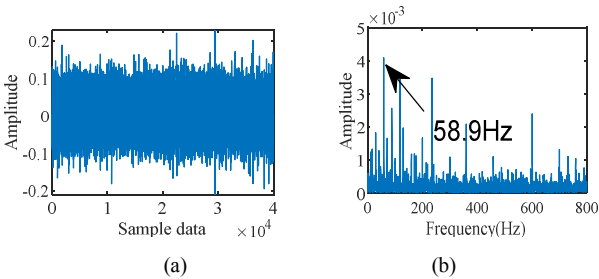
Fig. 7. The inner ring fault signal and its spectrum envelope spectrum. (a) The time-domain waveform of original signal. (b) The envelope spectrum of this signal.

The inner race signal is denoised by the fixed center frequency filter, which is a pre-processing.

Next the signal, which obtained from the first step, is respectively dealt by the SOGL and the FFT-OGS in frequency domain to get the new one. The noise is reduced largely, and the components of the new signal are more pure. Then the signal is transformed into time domain.

Finally, the real part of the signal is selected. The envelope spectrum is utilized to observe whether the important frequency is included in the reconstruct signal and can be found out well.

Comparing the results of envelope spectrum, the ability of the two methods in dealing with weak faults can be evaluated.
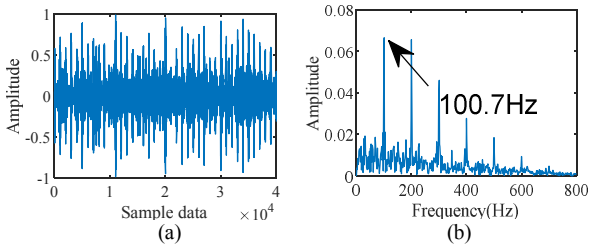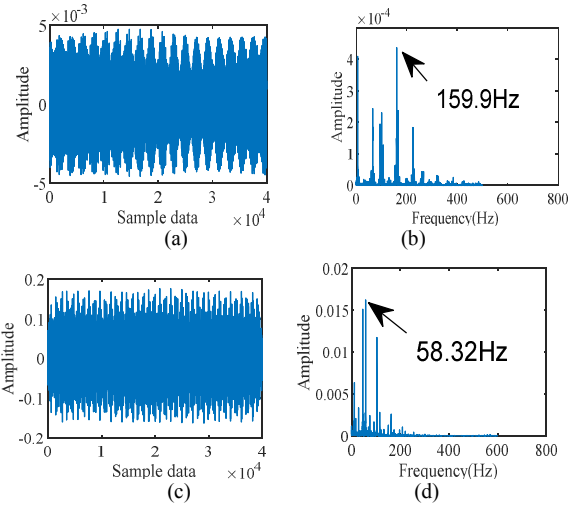

Fig. 8. The reconstruct signal and its spectrum envelope spectrum. (a) Reconstruct signal obtained from the method proposed in this paper. (b) The envelope spectrum of this reconstructed signal. (c)The signal obtained from FFT-OGL method. (d) The envelope spectrum of (c).

From the Fig. 8, the validity and superiority of the SOGL are more prominent. Although the FFT-OGS can reduce the noise, the characteristic frequency is also filtered out, which make the processing of signal is meaningless. For the bearing signal with weak fault, the purpose of signal processing is to highlight its characteristic frequency to help judge whether the fault exists and the type. As for the SOGL method proposed in this paper, it can highlight the vital frequency and reduce the interference of frequencies which occur due to noise.

V. CONCLUSION

In this paper, the SOGL which combined with l_1 norm and fixed center frequency filter is studied for early weak fault feature extraction of bearing. Experiments verify the superiority of this method. The conclusions are summarized as follows:

(1) The signal obtained from working rolling bearing is prone to be interfered with noise and occur coupling in time domain. Based on iterative shrinkage threshold (ISTA) and majorization–minimization (MM), the proposed algorithm processes signal in the Fourier domain to avoid the above problem, and can reconstruct the signal more accurately.

(2) The proposed method is applied to simulation signals and measured ones. From the experimental results, comparing with the FFT-OGS, the weak fault periodic pulse obtained by SOGL is obviously prominent.

REFERENCES

[1] R. Liu, B. Yang, E. Zio, X. Chen. "Artificial intelligence for fault diagnosis of rotating machinery: A review".Mechanical Systems & Signal Processing, vol. 108, pp. 33-47, 2018.

[2] R. Randall, J. Antoni. "Rolling element bearing diagnostics—A tutorial".Mechanical systems and signal processing, vol. 25, pp. 485-520, 2011.

[3] Z. Feng, M. Liang, F. Chu, "Recent advances in time–frequency analysis methods for machinery fault diagnosis: a review with application examples". Mechanical Systems & Signal Processing, vol. 38, pp.165–205, 2013.

[4] Y. Lei, J. Lin, Z. He, M. Zuo. "A review on empirical mode decomposition in fault diagnosis of rotating machinery". Mechanical Systems & Signal Processing, vol.35, pp. 108-126, 2013.

[5] M. Elad, M. Figueiredo. "On the Role of Sparse and Redundant Representations in Image Processing". Proceedings of the IEEE, vol. 98, pp. 972-982, 2010.

[6] K. Gröchenig. "Describing functions: Atomic decompositions versus frames". Monatshefte Für Mathematik, vol.112, pp.1-42,1991.

[7] S. Mallat, Z. Zhang. "Matching Pursuit with Time-Frequency Dictionaries". IEEE Transactions on Signal Processing, vol.41, pp. 3397-3415,1994.

[8] Y. ati, R. Rezaiifar, P. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition".Conference on Signals, Systems & Computers. IEEE, 2002.

[9] J. Trpp, A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit", IEEE Trans. Inf. Theory, vol 53, pp.4655– 4666, 2007.

[10] T. Rbert. "Regression shrinkage and selection via the lasso". Journal of the Royal Statal Society, vol 58, pp. 267-288,1996.

[11] M. uan, Y. Lin. "Model selection and estimation in regression with grouped variables". Journal of the Royal Statal Society,vol 68,pp.49-67, 2006.

[12] P. Cen, I. Selesnick. "Translation-Invariant Shrinkage/Thresholding of Group Sparse Signals". Signal Processing, vol 94, pp.476-489,2014.

[13] Z. Zhao, S. Wu, B. Qiao, S. Wang, X. Chen. "Enhanced Sparse Period-Group Lasso for Bearing Fault Diagnosis". Industrial Electronics IEEE Transactions on, vol 66, pp. 2143-2153, 2019.

A Novel Fault Diagnosis Method for Planetary Gearboxes under Imbalanced Data

Tianyu Gao
Department of Automatic Test and Control
Harbin Institute of Technology
Harbin, China
winser_hitgty@163.com

Jingli Yang*
Department of Automatic Test and Control
Harbin Institute of Technology
Harbin, China
jinglidg@hit.edu.cn

Shouda Jiang
Department of Automatic Test and Control
Harbin Institute of Technology
Harbin, China
jsd@hit.edu.cn

Abstract—To address the issue of fault diagnosis of planetary gearboxes under imbalanced data, a novel fault diagnosis method based on the improved energy-based generative adversarial network (IEBGAN) is proposed. Firstly, convolutional layers are added to the energy-based generative adversarial network (EBGAN) discriminator, thereby improving the feature extraction ability. Then, the classification loss is introduced into the loss function of EBGAN with the purpose of expanding the classification function of the discriminator. Finally, a planetary gearbox fault diagnosis model with sample generation capability is established to achieve the Nash equilibrium by the confrontation between the generator and the discriminator. Experimental results illustrate that the proposed method can improve the accuracy of fault diagnosis for planetary gearboxes even under imbalanced data.

Keywords—planetary gearboxes, fault diagnosis, imbalanced data, energy-based generative adversarial network

I. INTRODUCTION

With the advantages of compact structure, strong bearing capacity, smooth transmission and high efficiency, planetary gearboxes have become the core transmission components of large-scale mechanical equipment. However, the working conditions of these mechanical equipment are usually complex, and planetary gearboxes often run under dynamic and heavy load conditions, which lays hidden dangers for the faults of the key components [1]. In actual operation of mechanical equipment, it is easy to collect a large amount of monitoring data under normal conditions due to the stability of the planetary gearbox. In contrast, the monitoring data under fault conditions is hard to obtain, which leads to far more normal samples employed for the training of the fault diagnosis model than the fault samples. Under the condition of imbalanced data, there is much misdiagnosis for fault classes with fewer samples. Hence, it is significant to solve the issue of fault diagnosis under imbalanced data [2].

Generally speaking, the methods of fault diagnosis under imbalanced data are divided into algorithm optimization strategy and data synthesis strategy [3]. In the algorithm optimization strategy, the performance of the classifier is improved for the model training under imbalanced data. By adding weights representing the degree of imbalance in the loss function, Jia et al.[4] employed the deep normalized convolutional neural network (DNCNN) to adaptively process the classification of imbalanced data. To identify the fault classes of bearing with the imbalanced dataset, Duan et al. [5] introduced the binary tree to build a machine learning model based on the support vector data description (SVDD). Lan et al. [6] proposed a two-step framework to address the issue of imbalanced data, wherein the first step is based on weighted extreme learning machine (ELM) for anomaly detection, and the second step is based on preliminary ELM for fault

identification. By introducing the Laplace regularization term into the objective function of the deep autoencoder, Zhao et al. [7] constructed a fault diagnosis model with good generalization performance to perform feature learning and classification on imbalanced data. Zhao et al. [8] adopted normalized convolutional neural networks (CNN) to eliminate differences in the feature distribution. However, the algorithm optimization strategy is mainly aiming at the data of fixed imbalanced ratios, and lacking adaptability to various imbalanced ratios. The core idea of the data synthesis strategy is to construct a balanced data set by adding the data to minority classes, thereby obtaining a better classification result. Mao et al. [9] utilized the principle curve method to extract data distribution features, and employed the synthetic minority oversampling technique (SMOTE) to generate a normal dataset. On the basis of estimation for the imbalanced distribution of the radial basis function, Koziarski et al. [10] found the region suitable for generating the data of minority classes. Zhu et al. [11] divided sparse samples into inland samples, boundary samples and minority samples according to the location, thereby applying corresponding data generation strategies. Zhang et al. [12] proposed a weighted minority oversampling (WMO) method to balance the data distribution, where a clustering algorithm is employed to divide the sample subspace and adaptively synthesize samples. To adaptively extract effective features, the autoencoder is improved by adding the maximum entropy and sparse penalty to the loss function. To solve the issue of processing imbalanced data, Yang et al. [13] proposed an integrated diagnosis scheme based on fused autoencoder (FAE), which uses a variable-scale resampling strategy to compensate for information loss and skewed distribution. According to the sensitivity measurement (SM) standard, Babar et al. [14] designed an undersampling technique by using the multi-layer perceptron, which can preserve the distribution of sample information. The data synthesis strategy provides better adaptability to the imbalanced ratio, but its application effect depends on the quality of the synthesized data. Due to the complex distribution features of the monitoring data, it is difficult for the data synthesis strategy to learn the essential information and generate samples with high quality.

Generative adversarial network (GAN) is a generative model based on adversarial theory, which consists of a generator and a discriminator. The purpose of the former is to learn and imitate the distribution of real data, and the purpose of the latter is to accurately determine the data source. Both of them are optimized in the confrontation and eventually reach the Nash equilibrium. With the excellent ability in generating realistic appearance data, GAN has been successfully applied to plenty of fields. But the quality of the synthesized data is unstable because of the poor stability of the original GAN. Therefore, many researchers are devoted to improving the stability of the GAN model to obtain better quality synthetic

536

data. To solve the sample scarcity problem of the rolling bearing, Zheng et al. [15] introduced conditional GAN to generate multiple fault samples. Zheng [16] et al. designed a fault identification framework based on dual discriminator conditional GAN (D2CGAN), which helps to improve the quality of synthetic samples and avoid patterns collapse. Mao et al. [17] adopted GAN to supplement samples of minority classes, thus using the synthetic balanced data set to build a model based on the stacked denoising autoencoder (SDAE). Gao et al. [18] employed Wasserstein GAN with gradient penalty (WGAN-GP) to expand the fault samples of rolling bearings, wherein the gradient penalty term can improve the stability of the WGAN model. Liu et al. [19] developed a globally optimized GAN fault diagnosis approach for imbalanced data, where the DNN model is utilized as an additional discriminator for GAN. The main idea is alternately optimizing the generator and the two discriminators to simultaneously improve the sample generation quality of the GAN and the recognition ability of the model.

To address the issue of identifying the fault classes of the planetary gearbox under imbalanced data, a novel diagnosis approach based on improved energy-based generative adversarial network (IEBGAN) is proposed. The main contributions of this paper are as follows.

1. The EBGAN possesses the advantages of the stability of the adversarial training process and the diversity of generated samples. The convolutional layers are added to the autoencoder of the EBGAN discriminator to implement adaptive feature learning from the samples, thereby extracting the essential features of the samples.

2. The multi-classification function of the discriminator is realized by introducing classification loss.

3. The simulation experiments for fault diagnosis of the planetary gearbox are performed. Experimental results demonstrate that the approach can identify the fault classes of the planetary gearbox even under different degrees of data imbalances.

The paper is organized as follows. After the introduction, Section II illustrates the basic concept and procedures of the proposed method. To verify the performance of the method for fault diagnosis under imbalanced data, experiments are conducted in Section III. The conclusions are given in the last Section.

II. METHODOLOGY

A. EBGAN Model

To improve the stability of the GAN, Zhao et al. [20] introduced an energy-based model into the discriminator to construct the EBGAN model. In the EBGAN model, the discriminator is regarded as an energy function without an obvious probability explanation, which is a loss function that can be trained and optimized. The area close to the real data manifold is regarded as a low energy area, while the area far away is regarded as a high energy area. Similar to the original GAN, the generator will generate the minimum energy forged samples as much as possible during the entire confrontation training process, and the discriminator will try to give the fake samples high energy to determine these as false. By treating the discriminator as an energy function, a wider network structure and loss function can be adopted, instead of only a binary classifier with logistic output. According to the theory of EBGAN, Zhao et al. [20] proposed an implementation based on the structure of the autoencoder. The autoencoder reconstruction error in the discriminator is used as the energy

output, which improves the stability of the game process. The architecture of this EBGAN case is shown in Fig. 1, wherein *Enc* is the encoder in the discriminator, and *Dec* represents the decoder.
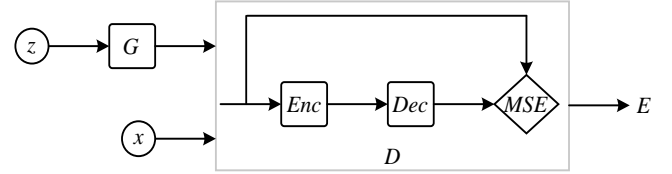
Fig. 1. Architecture of EBGAN

Given the positive margin m, the real sample x, the random noise z and generated sample $G(z)$, the objective function of the discriminator is expressed as

$$
\begin{aligned}
f_D(x,z) &= D(x) + \left[m - D(G(z)) \right]^+ \\
&= \left\| Dec(Enc(x)) - x \right\| \\
&\quad + \left[m - \left\| Dec(Enc(G(z))) - G(z) \right\| \right]^+
\end{aligned}
\tag{1}
$$

where $[\Box]^+ = \max(0,\Box)$, $Enc(\Box)$ represents encoding transformation, and $Dec(\Box)$ represents decoding transformation. The goal is that when the input samples are the real sample x, the value of $\left\| Dec(Enc(x)) - x \right\|$ is as small as possible, and when the samples are the generated samples, the value of $\left\| Dec(Enc(G(z))) - G(z) \right\|$ is as large as possible.

Similar to the original GAN, the objective function of the generator is as

$$
\begin{aligned}
f_G(z) &= \left\| D(G(z)) \right\| \\
&= \left\| Dec(Enc(G(z))) - G(z) \right\|
\end{aligned}
\tag{2}
$$

The goal of the generator is to make the value of $\left\| Dec(Enc(G(z))) - G(z) \right\|$ as large as possible, thereby deceiving the discriminator.

To ensure the diversity of the generated samples, a regular term is added into the objective function of the generator as follows.

$$
f_{PT}(S) = \frac{1}{bs(bs-1)} \sum_i \sum_{j \neq i} \left(\frac{S_i^T S_j}{\|S_i\| \|S_j\|} \right)^2
\tag{3}
$$

where $S = Enc(G(z))$ represents the energy expression of the generated samples, and bs is the batch size.

B. IEBGAN Model for Processing Imbalanced Data

To improve the feature extraction ability, the convolutional layers are introduced into the discriminator. In the encoder structure of the discriminator, two convolutional layers are adopted after an input layer with N neurons. The number of their convolution kernels is 32 and 16, and the size of the convolution kernels is 32 and 16, respectively. Relu is employed as the activation function of convolutional layers.

To improve the training speed and stability of the model, the batch normalization (BN) processing is performed on convolutional layers. Its operation is to normalize the output of convolutional layer and use the scaling and offset to process
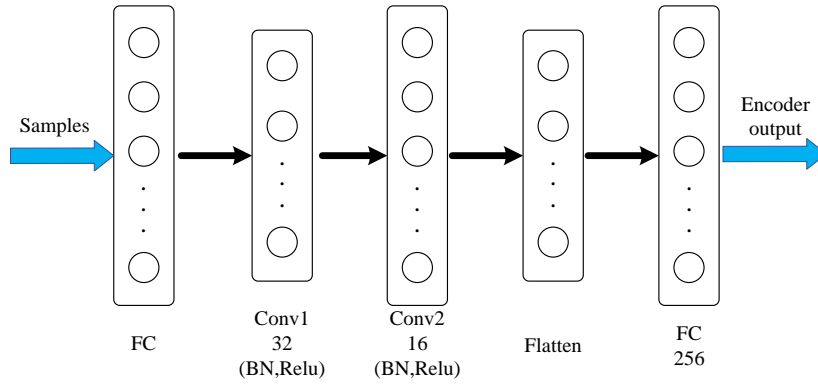
Fig. 2. Structure of the improved discriminator encoding network

the standardized amount with the purpose of enhancing the expressive ability. A fully connected layer with 256 neurons is used after the flatten layer. The structure of the improved discriminator encoding network is shown in Fig. 2.

To realize the sample classification under imbalanced data, the classification loss is introduced into the objective function of EBGAN, thereby expanding the multi-classification function of the discriminator. The classification loss f_c can measure the correctness of the output class, and its expression is as

$$f_c = E\left[\log P\left(C = c \mid x\right)\right] + E\left[\log P\left(C = c \mid G(z)\right)\right] \quad (4)$$

where $C \in \{1, 2, \ldots, n\}$, n is the number of sample classes, and $P(C = c \mid x)$ represents the probability the sample x belongs to class c.

Therefore, the objective function of the generator becomes $f_G + f_c$, and the objective function of the discriminator becomes $f_D + f_c$. Finally, IEBGAN model achieves the sample generation and classification by adversarial training.

C. Fault Diagnosis Process

The process of fault diagnosis method for the planetary gearbox based on IEBGAN is presented in Fig. 3. The EBGAN model adopted in this method helps to improve the quality and diversity of synthetic samples and avoid mode collapse. Convolutional layers are introduced into the EBGAN discriminator to enhance its feature extraction ability. Furthermore, the classification loss is added into the EBGAN loss function to realize the multi-classification function of the discriminator, thereby simplifying the fault diagnosis process. Therefore, with a good stability, the IEBGAN model can solve the problem of fault class identification for planetary gearboxes under imbalanced data. The specific steps of the fault diagnosis process are as follows.

1) Acquire vibration signals that can reflect the working state of the planetary gearbox.

2) Perform fast Fourier transform (FFT) on raw vibration signals to obtain the spectrum sample set.

3) Randomly divide the sample set into an imbalanced training sample set and a balanced test sample set.

4) Utilize the adversarial learning of the generator and discriminator to fine-tune the IEBGAN model parameters until meeting the Nash equilibrium.

5) Input the test sample set into the trained IEBGAN model to obtain the diagnosis results.

III. EXPERIMENTS AND RESULTS

To verify that the issue of fault diagnosis under imbalanced data can be addressed by the proposed method, Spectra Quest's Drivetrain Diagnostics Simulator (DDS) is adopted for data simulation. The fault diagnosis experiments are performed on a computer with 8GB RAM, and the IEBGAN model is built by using Python in TensorFlow environment.

A. Experiment Platform

The planetary gearbox experiment platform which is shown in Fig. 4 mainly consists of variable speed driving motor, two-stage planetary gearbox, two-stage parallel shaft gearbox, the programmable control panel, and the programmable magnetic controller.

As the power source of the transmission system, the variable speed driving motor with the power of 0~2.2kW and the speed of 0~5000r/min is responsible for generating driving torque. To simulate the operating environment of planetary
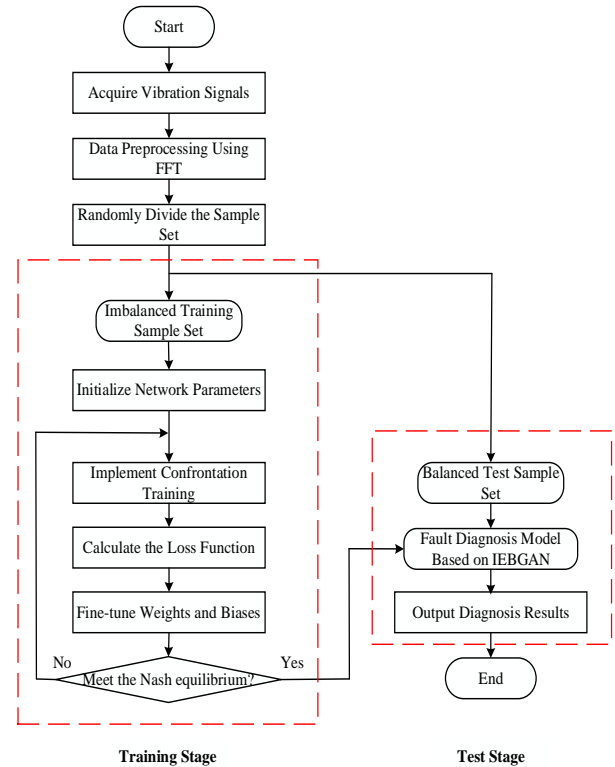
Fig. 3. Flow chart of the proposed fault diagnosis method
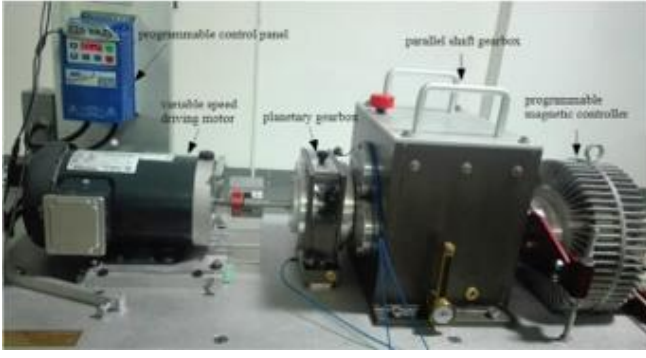
Fig. 4. Spectra Quest's Drivetrain Diagnostics Simulator

gearboxes under different working conditions, the control panel is used to adjust the speed, and the magnetic controller is adopted to adjust the load. The power transmission system consists of a transmission shaft, a parallel shaft gearbox and a planetary gearbox, which is mainly responsible for simulating the different fault classes of the transmission system.

B. Dataset

The sun gear of the planetary gear system is adopted as the main research object in this paper. Four typical fault classes are set in advance for the sun gear by wire cutting, which are missing tooth, cracked tooth, chipped tooth, and surface fault.

In the experiment, three driving motor speeds (1200r/min, 2400r/min and 3600r/min) and three load motor torques (0N • m, 20N • m and 40N • m) are set for the normal mode and the four fault modes, respectively, and nine operating conditions are formed by different speeds and different load combinations. Three piezoelectric acceleration sensors with model number 355B03 are selected to measure vibration signals at three positions on the outside, top and inside of the planetary gearbox. The German IMC SL-4 equipment is employed for data collection with the sampling frequency setting to 10kHz, so that the collected vibration signals contain 10,000 continuous data sampling points. The 120 samples of each state mode are collected under a working condition, thus, the number of samples in each state mode is 1080. Therefore, a total of 5,400 samples are collected in the normal mode and four fault modes of the planetary gearbox. For each fault class of the sun gear, 1,000 samples are randomly obtained to construct the experimental dataset. In the experimental dataset, 900 samples are randomly selected to build the training sample set for each fault class, and the remaining 100 samples are used as the test samples.

C. Experimental Parameter Settings

In the data preprocessing stage, FFT is performed on the original vibration signals to obtain the FFT spectrum signals. Because of carrying the valid information, the first 2000 points of the FFT spectrum signals are utilized as the input data of the fault diagnosis model.

The Adam optimizer with the learning rate of 0.001 is applied to the IEBGAN model, where the number of iterations is set to 1000. The margin distance in the loss function is set to 0.5.

In the actual operation of mechanical equipment, the normal samples of the sun gear are usually more than the fault samples due to the stability of the planetary gearbox. Therefore, the imbalanced data is simulated by setting the imbalanced ratios between normal samples and fault samples, where 6 imbalanced ratios are set from 100:50 to 100:1. Samples are drawn at random according to these imbalanced ratios to construct 6 imbalanced training sample sets and a balanced test sample set. All experimental data sets include 9 working conditions, where the setting is shown in Table I.

D. Results and Analysis

The accuracy of fault diagnosis is an overall evaluation of the performance of the classification model, which may ignore the diagnosis of minority class samples. To highlight the recognition effect, F-measure and G-mean are added for comprehensive evaluation in this paper. The two indicators can be calculated as

$$P = \frac{TP}{TP + FP} \tag{5}$$

$$R = \frac{TP}{TP + FN} \tag{6}$$

$$F - measure = 2 \times \frac{P \times R}{P + R} \tag{7}$$

$$G - mean = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}} \tag{8}$$

where *TP* is true positive, *TN* is true negative, *FP* is false positive, *FN* is false negative, *P* represents the precision rate, and *R* represents the recall rate. F-measure represents the harmonic average based on the precision and recall, and G-mean is used to maximize the geometric average of minority class accuracy and majority class accuracy. Both of them with

Table I EXPERIMENTAL SAMPLE SETS OF THE PLANETARY GEARBOX

| Dataset | State mode | | | | | Imbalanced ratio |
	Normal	Missing tooth	Cracked tooth	Chipped tooth	Surface fault	
Training set 1	900	450	450	450	450	100:50
Training set 2	900	180	180	180	180	100:20
Training set 3	900	90	90	90	90	100:10
Training set 4	900	45	45	45	45	100:5
Training set 5	900	18	18	18	18	100:2
Training set 6	900	9	9	9	9	100:1
Test set	100	100	100	100	100	1:1
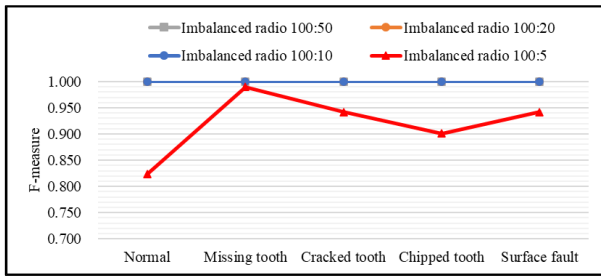
Fig. 5. F-measure of IEBGAN model under different imbalanced ratios
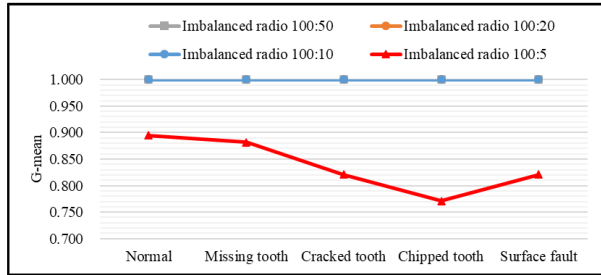
Fig. 6. G-mean of IEBGAN model under different imbalanced ratios

the range of [0,1] are evaluation indexes of the performance of the IEBGAN model under imbalanced data. The closeness of F-measure and G-mean to 1 indicates that their classifier has strong classification ability. F-measure and G-mean of IEBGAN model under different imbalanced ratios are shown in Fig. 5 and Fig.6, respectively.

It can be seen that the F-measure and G-mean of IEBGAN under various fault classes are equal to 1 until the imbalanced ratio is expanded to 100:5, and the cracked tooth and chipped tooth are greatly affected by the imbalanced data.

The IEBGAN model is compared with the CGAN model [15], D2CGAN model [16], GAN-SDAE model [17], WGAN-GP model [18] on fault diagnosis experiments for the planetary gearbox. The input data of the fault diagnosis model are data sets with various imbalanced ratio. The learning rate, batch size and number of iterations of all comparative models are the same as the proposed IEBGAN model. After many experiments, the diagnosis results of the proposed method and other methods for various imbalance ratios are shown in the Table II.

It can be found from Table II that the classification accuracy of all fault diagnosis models decreases with the increase of the imbalanced ratio. However, until the imbalanced ratio rises to 100:5, the IEBGAN model still has a classification accuracy of 94%, but the accuracy of other

methods has dropped significantly. When the imbalanced ratio rises to 100:2 and 100:1, the classification accuracy of all methods drops greatly and loses the classification ability. The diagnosis results demonstrate that the IEBGAN model can supplement effective samples and continuously enhance the classification performance under adversarial training, which solves the issue of insufficient fault samples to a certain extent.

IV. CONCLUSION

To improve the accuracy of fault class identification for the planetary gearbox under imbalanced data, a fault diagnosis method based on IEBGAN is proposed in this paper. First, the convolutional layers after batch normalization are added to the discriminator of EBGAN with the purpose of improving the ability of feature extraction. Then, the classification loss is introduced into the loss function of EBGAN, thereby realizing the fault class recognition function of the discriminator. Finally, by the confrontation training between the generator and the discriminator to achieve Nash equilibrium, a fault diagnosis model for planetary gearboxes with sample generation capability is established. The sun gear of the planetary gear is adopted as the main research object in this paper, and four typical fault classes are set in advance. Spectra Quest's Drivetrain Diagnostics Simulator is employed to collect the vibration signals of the planetary gearbox to construct an experimental data set, and the simulation experiment of fault diagnosis for the planetary gearbox is implemented. Experimental results show that the IEBGAN model can enhance the identification ability for planetary gearboxes even under imbalanced data compared with other typical fault diagnosis methods, and as the potential for improving the reliability and maintainability of planetary gearboxes.

REFERENCES

[1] C. Sun, and Y. Wang, "Advance in the study on fault diagnosis of helicopter planetary gears," Acta Aeronaut. Astronaut. Sin, vol. 38, pp. 020892, 2017.

[2] Y. Lei, F. Jia, and D. Kong, "Opportunities and challenges of machinery intelligent fault diagnosis in big data era," J. Mech. Eng, vol. 55, pp. 94-104, 2018.

[3] Y. Wang, G. Sun, and Q. Jin, "Imbalanced sample fault diagnosis of rotating machinery using conditional variational auto-encoder generative adversarial network," Applied Soft Computing Journal, vol. 92, pp. 106333, 2020.

[4] F. Jia, Y. Lei, and N. Lu, "Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization," Mechanical Systems and Signal Processing, vol. 110, pp. 349-367, 2018.

Table II CLASSIFICATION ACCURACY OF EACH METHOD AFTER EXPANDING FAULT SAMPLES UNDER VARIOUS IMBALANCED RATIOS

Model	100:50	100:20	100:10	100:5	100:2	100:1
CGAN[15]	99.80%	87.60%	78.20%	55.80%	44.80%	20.00%
D2CGAN[16]	99.40%	99.20%	99.00%	86.00%	32.00%	20.00%
GAN-SDAE[17]	100.00%	99.20%	97.40%	84.40%	59.60%	33.00%
WGAN-GP[18]	100.00%	94.40%	88.20%	64.40%	46.00%	25.80%
IEBGAN	100.00%	100.00%	100.00%	94.00%	49.40%	21.60%

[5] Y. Lan, X. Han, and W. Zong, "Two-step fault diagnosis framework for rolling element bearings with imbalanced data based on GSA-WELM and GSA-ELM," Proceedings of the Institution of Mechanical Engineers Part C Journal of Mechanical Engineering Science, pp. 095440621772809, 2017.

[6] X. Zhao, M. Jia, and M. Lin, "Deep Laplacian Auto-encoder and its Application into Imbalanced Fault Diagnosis of Rotating Machinery," Measurement, vol. 152, pp. 107320, 2020.

[7] B. Zhao, X. Zhang, and L. Hai, "Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions," Knowledge-Based Systems, vol. 199, pp. 105971, 2020.

[8] W. Mao, L. He, Y. Yan, and J. Wang, "Online sequential prediction of bearings imbalanced fault diagnosis by extreme learning machine," Mechanical Systems and Signal Processing, vol. 83, pp. 450-473, 2017.

[9] M. Koziarski, B. Krawczyk, and M. Wozniak, "Radial-Based oversampling for noisy imbalanced data classification," Neurocomputing, vol 343, pp. 19-33, 2019.

[10] T. Zhu, Y. Lin, and Y. Liu, "Improving interpolation-based oversampling for imbalanced data learning," Knowledge Based Systems, vol 187, 2019.

[11] Y. Zhang, X. Li, and L. Gao, "Imbalanced data fault diagnosis of rotating machinery using synthetic oversampling and feature learning," Journal of Manufacturing Systems, vol 48, pp. 34-50, 2018.

[12] J. Yang, G. Xie, and Y. Yang, "An improved ensemble fusion autoencoder model for fault diagnosis from imbalanced and incomplete data," Control Engineering Practice, vol. 98, pp. 104358, 2020.

[13] V. Babar, and R. Ade, "A Novel Approach for Handling Imbalanced Data in Medical Diagnosis using Undersampling Technique," Communications on Applied Electronics. Vol. 5, pp. 36-42, 2016.

[14] T. Zheng, L. Song, B. Guo, H. Liang, and L. Guo, "An Efficient Method Based on Conditional Generative Adversarial Networks for Imbalanced Fault Diagnosis of Rolling Bearing," Prognostics & System Health Management Conf. Qingdao, 2019.

[15] T. Zheng, L. Song, J. Wang, W. Teng, X. Xu, and C. Ma, "Data synthesis using dual discriminator conditional generative adversarial networks for imbalanced fault diagnosis of rolling bearings," Measurement, vol. 158, pp. 107741, 2020.

[16] W. Mao, Y. Liu, L. Ding and Y. Li, "Imbalanced Fault Diagnosis of Rolling Bearing Based on Generative Adversarial Network: A Comparative Study," IEEE Access, vol. 7, pp. 9515-9530, 2019.

[17] X. Gao, F. Deng, and X. Yue, "Data augmentation in fault diagnosis based on the Wasserstein generative adversarial network with gradient penalty," Neurocomputing, vol 396, pp. 487-494, 2020.

[18] F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, "Deep learning fault diagnosis method based on global optimization GAN for unbalanced data," Knowledge-Based Systems, vol 187, pp. 104837, 2020.

[19] J. Zhao, M. Mathieu, and Y. Lecun. "Energy-based Generative Adversarial Network," arXiv, 2016.

An Impedance-based Structural Health Monitoring by Using Piezoelectric Transducers

Xin Zhang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
zhangxin_m@yeah.net

Jiawen Xu*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
jiawen.xu@seu.edu.cn

Ruqiang Yan
School of Instrument Science and Engineering
Southeast University
Nanjing, China
yanruqiang@xjtu.edu.cn

Abstract—**Electromechanical impedance based (EMI) method is usually used health monitoring. This method requires accurately structural physical parameters and complex matrix operations. This paper presented an improved structural healthy monitoring method by using dual-piezoelectric transducers. The dual-piezoelectric transducers which have one transducer as actuator and another one as sensor were utilized to decouple electrical impedance of measured structure. Convolutional neural network (CNN) model was used to classify response signal from different structures with high accuracy.**

Keywords—Piezoelectric transducer, Impedance-based method, Convolutional neural network

I. INTRODUCTION

Early diagnosis identification, especially the detection and identification of minor faults, has an important role in the field of structural healthy monitoring (SHM). To detect locations and severity of structural damage, extensive works have been done [1]. Among which EMI methods have superiorities of simple data acquisition, ability to present global structural information, and independent analytical models[3-4]. Furthermore, the EMI method enjoys applicability of structures which is not only made of metallic but also be made of ceramic and composite [2, 3, 5,6].

In a typical EMI method measuring scenario, the mechanical impedance of interested system couplings with electrical-impedance of a piece of lead zirconate titanate (PZT) transducer which is attached on the structure[7]. The fact has been demonstrated that electrical impedance frequency response function (FRF) shifting of PZT associate with changes of properties in measuring system. Therefore, the electrical impedance shifting of PZT can be utilized to identify properties change of system[2,3]. When a reliable model of structure is available, the electromechanical impedance/ admittance curve of measured structure can be used to inverse calculate and identify the location, category and severity of structural faults [10]. However, this method requires accurately structural physical parameters and complex matrix manipulation. The effectiveness of fault diagnosis results rely on the accuracy of modeling of the structure under test. In addition, the noise in measurement will deteriorate the results of this method.

Neural networks had been applied to the field of structural healthy monitoring [9][11][12]. For example, Oliveira has explored the effectiveness of CNN model which was applied to structural healthy monitoring in aluminum structures which revealed an effective pattern classification for SHM approaches [13].

In this research, we utilize dual-transducers which have one PZT as a senor and the other one as an actuator to acquire structural electrical impedance. Comparing to the experimental model with one piece of PZT, the dual-transducer schema can decouple structural impedance directly from the response of sensor PZT by introducing phase sensitive detection(PSD) algorithm. PSD algorithm is a kind of signal detection approach which can extract the amplitude of response signals under excitation frequencies without complex calculation of matrix operation. A CNN model is then adopted to classify response signals from different structures.

The rest of this paper is organized as follows. The illustration of physics of the proposed schema is arranged in section 2. In section 3, experimental studies and analysis of results are presented, including description of PSD signal processing algorithm and discussion of classification results by CNN. Conclusion is given in section 4.

II. CONFIGURATION AND MODELNG

A. System Illustration
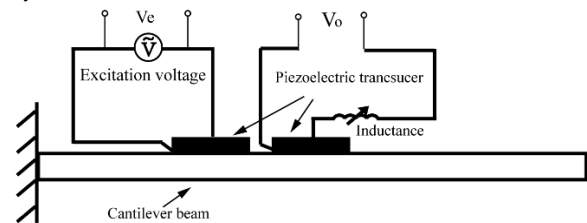
Fig. 1. Dual-transducers EMI health monitoring system.

To solve the problem that mechanical impedance couples with electrical impedance of PZT, dual-transducers scenario is presented in Fig. 1. Two transducers are attached on the surface of composite cantilever. While frequency-sweeping excitation voltage is applied to one of PZT transducers, response signals of another transducer can be utilized to extract FRF of cantilever beam. The tunable inductance is in series with sensor PZT in the circuit. Consequently, the system dynamics response of cantilever beam can be detected by measuring voltage of the second PZT directly without decoupling. The acquired signals can be processed by PSD algorithm in following steps. In next section, finite element method (FEM) and the underlying discussing are presented which can illustrate the effectiveness of proposed dual-transducers configuration.

B. FEM Model

The piezoelectric composite cantilever which shown in Fig. 1. is discussed by a FEM model in equation (1). The

electromechanical coupling equations of proposed dual-piezoelectric transducer scenario are given,

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} + \mathbf{k_{121}}Q_1 + \mathbf{k_{122}}Q_2 = 0 \qquad (1a)$$

$$\frac{1}{C_p}Q_1 + \mathbf{k_{121}^T}\mathbf{q} = \tilde{V}_e, \; L\ddot{Q}_2 + \frac{1}{C_p}Q_2 + \mathbf{k_{122}^T}\mathbf{q} = 0 \qquad (1b)$$

where, $\mathbf{K}$, $\mathbf{C}$ and $\mathbf{M}$ is the matrix of the stiffness, damping and mass of the cantilever system. The electromechanical coupling matrixes of the two PZTs are $\mathbf{k_{121}}$ and $\mathbf{k_{122}}$ respectively. Besides, C_p is the capacitance of the identical PZTs. The displacement vector of the system is expressed as $\mathbf{q}$. Charges on the electrodes of piezoelectric transducers are represented by Q_1 and Q_2, respectively. L is the inductance load connected to the secondary transducer. $\tilde{V}_e$ expresses excitation voltages applied on one of PZTs which named actuator PZT. In present system, the output voltages on the sensing PZT is calculated

$$Z_{dual} = \frac{\tilde{V}_{out}}{\tilde{V}_e} = \frac{\mathbf{k_{122}^T}\mathbf{k_{121}}C_p}{-\omega^2\mathbf{M}+i\omega\mathbf{C}+\mathbf{K}} \qquad (2)$$

where ω is the frequency of excitations. It is obvious in equation (2) that the calculated impedance in transducers is inversely proportional to the impedance term of the cantilever, the $-\omega^2\mathbf{M}+i\omega\mathbf{C}+\mathbf{K}$. Besides, it indicates that the response signal getting from the sensing PZT only contains structural dynamics of the cantilever. Namely, It is successfully to decoupling mechanical impedance of cantilever from the capacitance impedance of the piezoelectric transducer. Piezoelectric transducers usually contains high capacitance impedance. In other words, it is difficult to decouple and amplify the structural impedance because the capacitance impedance of piezoelectric transducers dominate responses. Hence, the structural mechanical impedance is decoupling from the electrical one in the new configuration. It is possible to utilized this configuration in further signal processes.

C. Signal Processing Procedure
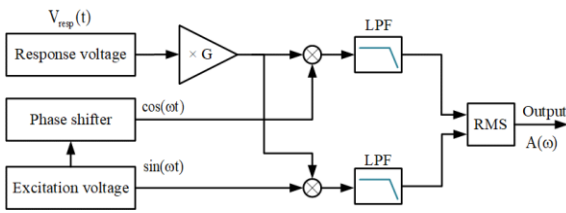
Fig. 2.   Data processing flow

Most signal processing procedure in EMI methods health analyze the raw response signal from the piezoelectric transducer. Statistical indices such as the resonant frequencies shifting was further utilized to achieve damage detection. Consider that the measured impedance signals would be extremely weak which might be merged into the noises. Hence, the PSD method is introduced to extract the measured signal which can represent the mechanical impedance of piezoelectric composite cantilever in this paper. The explanation of implementing the PSD into SHM system configured by dual-transducer and EMI method is explained as following.

The procedure of response and the excitation signal in PSD method processing is presented in Fig. 2. Firstly, the sweeping frequency sine voltage excitations are applied to the PZT which acts as an actuator. Because the cantilever can be modeled as a linear time-invariant system, response signals have the same frequencies as excitations. The response signal of the cantilever system is collected from another sensing PZT.

Fig. 2. provides the details for the procedure of the cross-correlation signals processing and analyzes the principle of noise reduction in the PSD method. The response is given

$$V_{resp}(t) = A\sin(\omega t + \alpha) + n(t) \qquad (3)$$

where A is amplitude of response signals getting from the sensing PZT, ω is frequency of excitations. $n(t)$ stands for noises and denotes the phase. In this case, multiplying $V_{resp}(t)$ by two reference signals orthogonal with each other, $\sin(\omega t)$ and $\cos(\omega t)$, yields.

$$V_{resp}(t)\times\sin(\omega t) = \frac{1}{2}A\cos\alpha - \frac{1}{2}\cos(2\omega t+\alpha) + n(t)\sin(\omega t) \quad (4a)$$

$$V_{resp}(t)\times\cos(\omega t) = \frac{1}{2}A\sin\alpha + \frac{1}{2}\sin(2\omega t+\alpha) + n(t)\cos(\omega t) \quad (4b)$$

where prosthaphaeresis formulas are adopted in the derivations of (4). The result is obtained that productions of signals yield two DC terms $\frac{1}{2}A\cos\alpha$ and $\frac{1}{2}A\sin\alpha$. The two terms are associating with the amplitude and phase of responses which realize decouple mechanical impedance of the cantilever from the electromechanical impedance of PZT. Furthermore, works have done to filter the resulting signals after arithmetic products between reference signals and amplified signals by low-pass filters respectively. The result is that the quadratic sum of the two DC terms, $\frac{1}{2}A\cos\alpha$ and $\frac{1}{2}A\sin\alpha$, would contain the structural impedance feature associating with amplitude of responses at excitation frequency ω only.

D. Structural Healthy Monitoring Using 1D-CNN
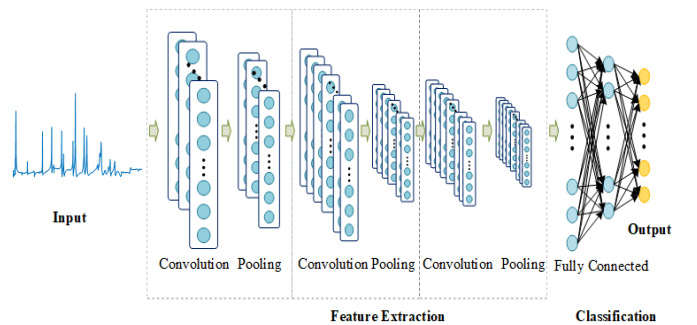
Fig. 3.   Architecture of the 1D-CNN

CNN is a deep linear network that has achieved success in computer vision. One-dimensional CNN (1D-CNN) has shown good performance in time series processing. In this research, the final signal processed by PSD is a one-dimensional signal which can present the frequency response

function of the system. Hence, we utilized 1D-CNN to classify data of different structures.

Fig. 3. shows the architecture of the 1D-CNN adopted for our structural impedance signal process. Convolution neural network generally includes three structures: convolution layer, pooling layer, and full connection layer. Generally, one or more convolution layers will be added to form a convolution module. Several similar convolution modules are superimposed to form a deep convolution network. Finally, full connection layers are added to realize classification or regression. Different classification results correspond to different structural faults.

III. RESULT AND DISCUSSION

A. Experimental Set-up
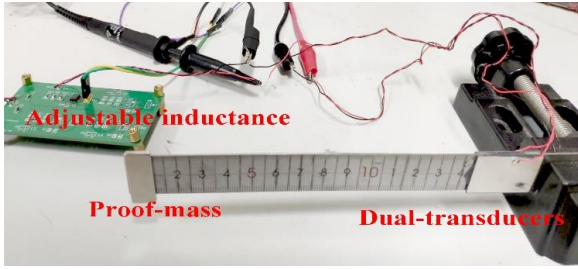
Fig. 4.   Experimental configuration

Experiments by using dual PZT patches are performed to validate the EMI methods in SHM. The experimental configuration is shown in Fig. 4. The cantilever is mounted onto a fixture. Dual-transducers are pasted on the two sides of the experimental stainless steel beam. The two identical PZTs are bonded on the surface of cantilever in vicinity of fixture. The mass block was bonded at another end of the cantilever. The dynamic response data of the system can be enhanced by a variation of the inductance of the peripheral circuit [8]. In the experiment, frequency sweeping sinusoidal excitations were applied to actuator PZT. The amplitude of excitation voltage is 3V. The acquisition system was developed in SIMULINK and obtain responses from PZT. Besides, the amplitude of responses extracts algorithm PSD under excitations was implemented in the SIMULINK. Forward sweeping frequency ranged from 50Hz to 2550 Hz with a step of 10Hz/s.

B. Experimental Results

In this section, we obtained responses from the various structural conditions to evaluate the proposed method. Besides, 1D-CNN was utilized to classify structural conditions. A small mass block was attached to the cantilever. The structural conditions vary by different locations of the mass block. The mass block locates at the end of the cantilever firstly. The location of the mass block changes 0.5mm when moving it. Locations and number of samples are shown in Table I of five kinds of conditions.

TABLE I.   NUMBEER OF SAMPLES

Condition	L1	L2	L3	L4	L5
Location(mm)	0.0	0.5	1.0	1.5	2.0
Number of samples	61	66	61	61	61

As shown in Fig. 5. ,responses of three cases are presented. It can be obtained that response signals have peaks at the resonant frequency. Meanwhile, the response signals of frequency shift between different structures. Hence, classification methods can be adopted to realize structure health monitoring.
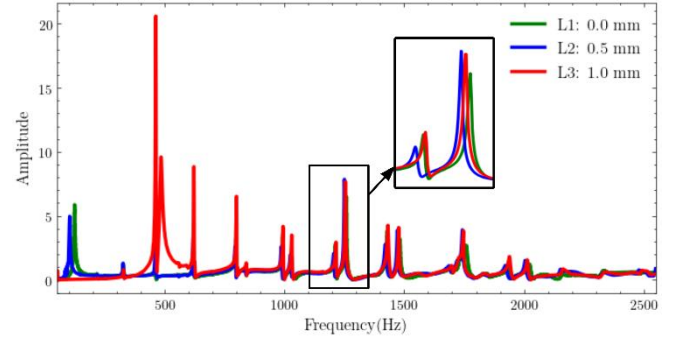
Fig. 5.   Response signal with different location of mass block

C. Classification Results

1D-CNN was adopted to extract features of the structure. The input of the network is a time series of signals processed by PSD. The output of the network is a probability of a type of structural condition. The first, second, and third are convolutional layers which have Rectified Linear Unit (ReLU) as activation functions. The three convolution layers are composed of 16,32 and 32 filters of length 3. The last of the networks are fully-connected layers. The last output layer is fed to a softmax classification layer. This network trains by a loss function of categorical cross-entropy, optimizer of Adam and batch size of 16. The final accuracy achieves 100%. Fig. 6. shows the confusion matrix of training results. It can be seen that each structure has been distinguished.
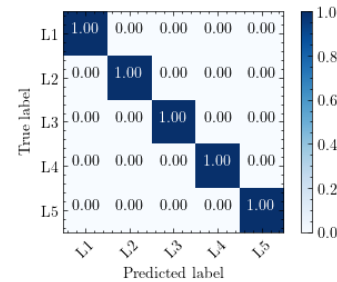
Fig. 6.   Confusion matrix

IV. CONCLUSIONS

In conclusion, a configuration based on dual-transducer for structural impedance measuring was proposed in this paper. The conceptual illustration adopts piezoelectric cantilever which was modeled by FEM. Experiments have been fabricated with dual PZTs configuration pasted on a stainless steel beam. Based on the underlying model analysis and experimental results, it can be concluded that the impedance FRF of the structure can be calculated from responses of PZT directly by PSD method. Finally, the effectiveness of using 1D-CNN to classify different structures experimentally. It is potential to apply this dual-transducer based configuration in SHM without knowing the exact parameters of structure.

References

[1] J. Yu, H. Zhu, and M. Huang, "Numerical study of structure health monitoring using surface-bonded and embedded PZT transducers," 2010 Int. Conf. Mech. Autom. Control Eng. MACE2010, no. 50778077, pp. 895–898, 2010.

[2] W. S. Na and J. Baek, "A review of the piezoelectric electromechanical impedance based structural health monitoring technique for engineering structures," Sensors (Switzerland), vol. 18, no. 5, 2018.

[3] X. Qing, W. Li, Y. Wang, and H. Sun, "Piezoelectric transducer-based structural health monitoring for aircraft applications," Sensors (Switzerland), vol. 19, no. 3, pp. 1–27, 2019.

[4] W. Fan and P. Qiao, "Vibration-based damage identification methods: A review and comparative study," Struct. Heal. Monit., vol. 10, no. 1, pp. 83–111, 2011.

[5] C. Zuo, X. Feng, Y. Zhang, L. Lu, and J. Zhou, "Crack detection in pipelines using multiple electromechanical impedance sensors," Smart Mater. Struct., vol. 26, no. 10, 2017.

[6] P. Negi, T. Chakraborty, N. Kaur, and S. Bhalla, "Investigations on effectiveness of embedded PZT patches at varying orientations for monitoring concrete hydration using EMI technique," Constr. Build. Mater., vol. 169, pp. 489–498, 2018.

[7] C. Liang, F. P. Sun, and C. A. Rogers, "Coupled electromechanical analysis of piezoelectric ceramic actuator-driven systems: determination of the actuator power consumption and system energy transfer," Smart Struct. Mater. 1993 Smart Struct. Intell. Syst., vol. 1917, no. September 1993, pp. 286–298, 1993.

[8] J. Kim and K. W. Wang, "An enhanced impedance-based damage identification method using adaptive piezoelectric circuitry," Smart Mater. Struct., vol. 23, no. 9, 2014.

[9] J. Xu, J. Dong, H. Li, C. Zhang, and S. C. Ho, "Looseness monitoring of bolted spherical joint connection using electro-mechanical impedance technique and BP neural networks," Sensors (Switzerland), vol. 19, no. 8, 2019.

[10] A. A. Ezzat, J. Tang, and Y. Ding, "A model-based calibration approach for structural fault diagnosis using piezoelectric impedance measurements and a finite element model," Struct. Heal. Monit., 2020.

[11] Z. Wu, H. Jiang, K. Zhao, and X. Li, "An adaptive deep transfer learning method for bearing fault diagnosis," Measurement, vol. 151, p. 107227, 2020.

[12] C.-U. Nguyen, T.-C. Huynh, and J.-T. Kim, "Vibration-based damage detection in wind turbine towers using artificial neural networks," Struct. Monit. Maint., vol. 5, no. 4, pp. 507–519, 2018.

[13] M. A. de Oliveira, A. V. Monteiro, and J. V. Filho, "A new structural health monitoring strategy based on PZT sensors and convolutional neural network," Sensors (Switzerland), vol. 18, no. 9, 2018.

Tip Timing Based Non-contact Vibration Measurement of Aero-engine Turbine Blades

Meiru Liu
The State Key Laboratory for
Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
18784010242@163.com

Weiqiang Gao
Sichuan Gas Turbine Establishment
Aero Engine Corporation of China
Mianyang, China
yishi9566@qq.com

Xiao Xiao
Sichuan Gas Turbine Establishment
Aero Engine Corporation of China
Mianyang, China
xiao-ganr@126.com

Guangrong Teng
Sichuan Gas Turbine Establishment
Aero Engine Corporation of China
Mianyang, China
492722046@qq.com

Chunyan Ao
The State Key Laboratory for
Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
aochuanyan7@stu.xjtu.edu.cn

Baijie Qiao*
The State Key Laboratory for
Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
qiao1224@xjtu.edu.cn

Abstract—**Non-contact vibration measurement based on tip timing has been widely used in aero-engine blade health monitoring. However, it is different to inspect the vibration of turbine blades with small size in high rotating speeds, temperature and pressure operation. To overcome the difficulty of aero-engine turbine blade measurement, firstly, the optical sensor with cooling channel is designed and a novel air-cooling method is proposed. Secondly, a disk for once-per-revolution (OPR) sensor placement is designed and the black paint is sprayed on the shaft. The OPR signal can be achieved by detecting the intensity change of light as the shaft rotates. Thirdly, the non-contact vibration measurement based on tip timing is conducted to monitor turbine blades, and the resonance frequencies, blade tip vibration amplitude and rotational speeds of blades are obtained. The test results show that the 1st bending mode resonance of the turbine blades occurred at about 8200r/min excited by engine order 12 and 9100r/min excited by engine order 11. Furthermore, the identified vibration parameters from the test have a good match with the results of the numerical calculation. It demonstrates the promise of the proposed non-contact vibration measurement technique based on blade tip timing for health monitoring of aero-engine turbine blades.**

Keywords—aero-engine turbine blade, non-contact vibration measurement, tip-timing, cooling channel, once-per-revolution signal

I. Introduction

Turbine blades are increasingly confronted the operation condition of high temperature, high rotational speed and severe loading, and prone to the unexpected and catastrophic failure. It raises great concern with respect to safety and optimal outage planning and demand for enhanced Remaining Useful Life (RUL) estimation[1-3].

The traditional strain measurement method uses strain gage (SG) attached on the blade and requires to measure the dynamic strain in the operation of high temperature and speed. However, the survival rate of SG is usually very low under the high temperature air flow in engine, which causes the failure of SG measurement[4],[5]. Therefore, the dynamic strain acquisition through SG is far from ideal as their low survival rate and intrusive attributes for measurement. Non-contact measurement based on blade tip timing (BTT) has been developed for many decades as an attractive vibration monitoring technique for turbine and compressor blades. BTT technique makes it possible to detect the vibration state of all the blades on the same stage of compressor and turbine. This is achieved by using a set of proximity BTT probes placed circumferentially around the casing which can record the different arrival time of each passing blade resulting from vibration[6]. The obvious advantages of BTT are the non-intrusive attributes and availability for online monitoring. For the above virtues, the non-contact vibration measurement based on tip timing is regarded as the most promising technique to replace strain gage to monitor the health of rotor blades, thus preventing unexpected and catastrophic failures[7].

The non-contact vibration measurement based on tip timing has been widely used for rotor blades health monitoring and failure inspection. Du Toit[8] developed a stochastic hybrid blade tip timing approach to identify and classify turbomachine blade damage considering the measurement uncertainty. Mohamed[9] used a novel method for the determination of the change in blade tip timing probe sensing position due to steady movements of blades, which greatly improved the accuracy of vibration acquisition. Ouyang et al.[10] developed a method for analysing the synchronous and nonsynchronous vibration of blades by non-intrusive measurement and designed a hardware system based on blade tip timing technique. Liu[11],[12] detected the vibration characteristic of compressor blades and identify the cracking failure based on blade tip timing measurement. In aforementioned work, blade tip timing technique is usually applied on the compressor monitoring. However, there are still a few problems on the application of BTT technique on turbine rotor blades monitoring. First, the operating temperature of the turbine blades are much higher than the limited temperature of BTT optical sensor. Second, the rotational speed signal from OPR sensor as the vital part of the vibration analysis is hard to achieve accurately. Finally, the vibration magnitude of turbine blades is small, which rises the acquirement for the sensitivity, measurement accuracy and available data processing.

In this paper, for addressing the above three key problems in BTT measurement, first, a cooling channel is designed on the optical sensor and the air cooling method is adopted to cool down the sensor to below 500°C. Second, the OPR signal is achieved through the sensor installed on a circle blisk by detecting the different light density within each revolution. Third, a novel arrangement for BTT optical sensors on the casing is designed to measure the weak vibration of turbine blades with small amplitude. Finally, the non-contact vibration measurement technology is performed to measure

---

* Baijie Qiao is the corresponding author. (e-mail: qiao1224@xjtu.edu.cn).

the turbine blade vibration and the key resonance parameters of all the blades on the same stage are obtained. The experimental results are compared to those by the numerical calculation to demonstrate the availability of the proposed measurement technique for aero-engine blade monitoring based on BTT.

II. THEORY OF NON-CONTACT VIBRATION MEASUREMENT BASED ON BLADE TIP TIMING

A. Basic Theory

The principles of non-contact vibration measurement based on BTT are shown in Fig. 1. The BTT sensors are installed on the casing to record the time of arriving (TOA) of each passing blade. Assuming that there is no vibration when blades rotate, the TOA of each blade on the rotor is uniform and can be calculated by the rotational speed and the relative position on the rotor. However, when the blade vibrates, the deflection of blades can cause the early or late TOA. Hence, the vibration can be calculated through the difference of TOA referring to the standard TOA without deflection in conjunction with the rotational speed signal from OPR sensor. Furthermore, the blade key parameters such as vibration amplitudes and frequency can be obtained by some specific data processing algorithms[13].

B. Data Processing Method

Because of the usual high vibration frequency of blade and limited measurement positions for BTT sensor, the captured vibration signal by BTT is seriously under-sampled which does not satisfy the Shannon sampling theorem. Therefore, the traditional methods based on Fourier transformation are not available for BTT data.

Single degree of freedom (SDOF) and circumferential Fourier fitting (CFF) are two classical methods widely used in BTT data analysis, which can identify the synchronous vibration parameters of blades from the under-sampled signal. The same assumption for the two methods is that the blade vibration has a sinusoidal form. The difference of the two methods is that SDOF is available for the blades in the acceleration or deceleration run while CFF is effective for signal samples at the constant rotational speed[14].
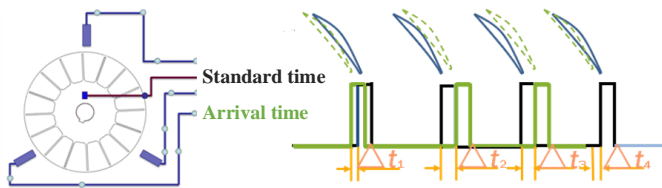
Fig. 1 Principle of non-contact vibration measurement system

The principle of SDOF is that the BTT sensors detect the vibration in every revolution when blades rotate. As the speed increases or decreases, the sensors sample the discrete vibration data when blades pass the resonance region. All the data are combined to fit the blade vibration curve which includes the information of vibration amplitude, resonance speed and damping. Different curves fitted by SDOF are varied in vibration phase. The resonance frequency can also be calculated with the prior knowledge of engine order (EO)[15]. SDOF can be used in speed fluctuation measurement and usually one or two sensors can achieve the vibration inspection. However, SDOF is insufficient in amplitude accurate identification.

For CFF, at least three BTT sensors are needed to perform the measurement. Based on the least square algorithm, the vibration curves of blades can be fitted by using the discrete sampled data from all the sensors. The remarkable advantage of the method is the ability for couple-modes vibration analysis of blades with relatively high accuracy, which can provide abundant vibration parameters from the fitted curves such as amplitude, phase, damping and fitting errors. However, the sensor arrangement on the casing has great effect on the parameter identification accuracy.

III. NON-CONTACT MEASUREMENT CONFIGURATION OF TURBINE BLADES

A. Optical Sensor with Air Cooling Channel

Non-contact vibration measurement for turbine blades is performed by using the optical sensor designed with one cooling channel, as shown in Fig. 2. The exact size and shape of the sensor are well designed and compact. The total length of the sensor is roughly 1.8m which is divided into three parts as the head, middle part and tail. The head is mounted on the casing of the turbine. The head has a special air cooling channel and a thermocouple. The air cooling channel is adopted to cool the head and the thermocouple is to measure the operation temperature of the head. The thermocouple is conducted at the end of the head while the light fiber transmission, light fiber receiving and air cooling by inducted tube are achieved at the middle part with probable length of 3.2cm.
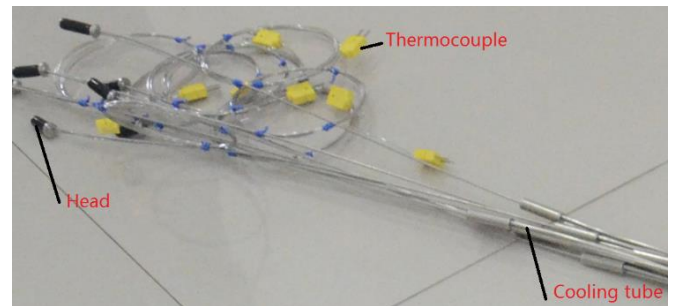
Fig. 2 Optical sensor with air cooling channel

B. Cooling Method of the Optical Sensor

The maximum work temperature of the optical sensor with sapphire fiber is roughly 650°C without cooling. In order to keep the measurement performance in the temperature range from 900°C to 1200°C, the sensor must be designed by the suitable cooling channel.

In general, the water and air are the most used as the cooling medium. As the turbine blisk is very thick and relatively short in axial direction, it is hard to add the water flowing facility. Meanwhile, it is prone to generate failure as the water can easily leak. Therefore, the air cooling method is more appropriate in practice. There are two air cooling patterns. The first one is to adopt air from the relative test equipment and the applying pressure of the air is greater than that of the inner flow in the turbine when the engine is operating. Through this pattern, the outlet air can be put into the inner flow channel of the turbine to cool the sensor. Another method is to fix the sensor installation position closing to the flowing air directly from the last stage of compressor. The maximum temperature of the air is about 500°C which is capable to cool the sensor down to the promising temperature. In this paper, the above two methods are used

simultaneously to ensure the designed BTT sensors can be cooled to the acceptable temperature.

C. Sensor Arrangement

The sensor arrangement has great significance on the measurement accuracy of BTT system[16],[17]. For achieving the effective non-contact measurement, it needs one sensor on the chord length direction and 3~5 sensors circumferentially mounted on the casing. Hence, the vibration parameters such as resonance frequency, amplitude and speed can be calculated by subsequent data processing. In order to inspect the maximum blade deflection along circumference, positions near the leading or trailing edge are promising but also prone to miss the signal when the sensor is amounted too close to the edge as the axial movement occurs on the blade. In such a case, the BTT measurement cannot be executed effectively. Before the sensor arrangement performed, the axial movement of rotating blades should be considered. Provided the effective measurement of sensors, the mounted position should be as close as possible to the leading or trailing edge of the blade to capture the prominent deflection by BTT sensor. In this test, combining the above considerations, the position with 4.3mm distance closing to the leading edge is determined for the axial installation of BTT sensors as shown in Fig. 3.
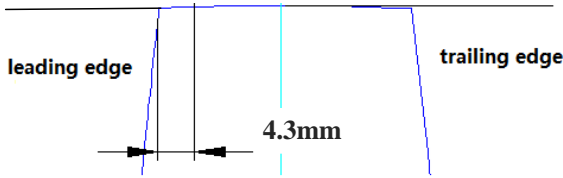
Fig. 3 Axial location of the sensors
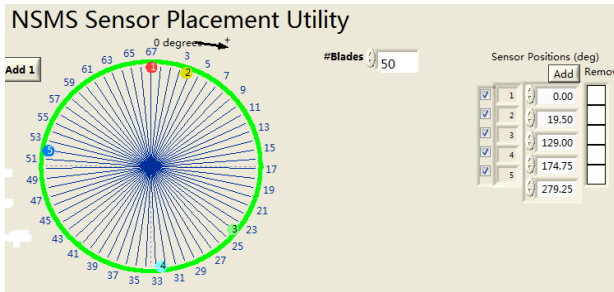
Fig. 4 Circumferential location of the sensors

For the circumferential arrangement of BTT sensors, the CFF method acquires at least 3 sensors embedded on the casing to monitor the vibration of whole blades on the same stage. In this study, 5 sensors are mounted circumferentially on the casing for multi-EO identification and accurate measurement. In this test, some special positions like those for the oil tube installation should be excluded. Performing the arrangement optimization considering 8 EO, 9 EO, 10 EO, 11 EO, 12 EO and 45 EO, the installation angles of 5 sensors are set as 0°, 19.5°, 129°, 174.75° and 279.25° as shown in Fig. 4.

D. The Achivement of Once-per-revolution Speed Signa

The accurate OPR speed signal is the foundation of the effective vibration recovery by non-contact measurement. In general, two ways can achieve the OPR speed signal. The first one requires a keyway on the shaft, which can be detected by the magneto electric tachometric transducer at each revolution. The another method is to make use of the optical probe sensing the light intensity changing provided that a black paint is

intentionally sprayed onto the shaft. In the BTT rotor test, the performance of two methods is determined based on the structure of the engine itself.

For the engine used in the test, it is difficult to install multiple fiber sensors on the casing. Moreover, the fiber is prone to be damaged and fail to get effective measurement. Therefore, a single disk is designed to rotate with the shaft and the magneto electric tachometric transducer is fastened on it to measure the OPR signal efficiently.

IV. FINITE ELEMENT ANALYSIS

In order to validate the test results, the finite element analysis is applied in ANSYS. The finite element (FE) model is meshed by SOLID 92 element with 10 nodes. The blade FE model has 425317 elements and 668785 nodes in total.

For the numerical calculation in ANSYS, the blade FE model is loaded by the aerodynamic force, temperature and rotational speed. Different boundary condition should be applied on the model when various numerical calculations are performed. For strength analysis, the normal displacement of all the compressive plane of the blade tenon is constrained and the axial displacement of three points on the inlet edge of the tenon is constrained. The numerical results are obtained by the blade FE model. The static natural frequency of 1ST bending mode of the turbine blades is 1722.3Hz and the corresponding dynamic frequency at the maximum operation speed (13000r/min) is 1602.0Hz.

V. RESULTS OF NON-CONTACT VIBRATION MEASUREMENT

A. 1ST Bending Resonance of 12EO

In the test, the engine firstly starts up to a relatively low speed of 6960r/min. Then the shaft ramps up to the maximum speed of 13000r/min with the acceleration of 25(r/min)/s. After staying the operation of the maximum speed for 3min, the engine decelerates to the stalling with the same changing speed. The non-contact measurement system monitors the turbine blades vibration and the signals sampled by optical sensors are pretty well.

The results of the non-contact measurement system show that all the turbine blades resonate at the speed of 8500r/min roughly. With SDOF and CFF algorithms, the important vibration parameters like resonant speeds, frequencies and amplitudes of all the blades can be determined. The results of the 27th blade analyzed by SDOF and CFF are respectively shown in Fig. 5 and Fig. 6.

Fig. 5 SDOF result of 27th blade from 1st sensor at the rotational speed of 8043.2 r/min
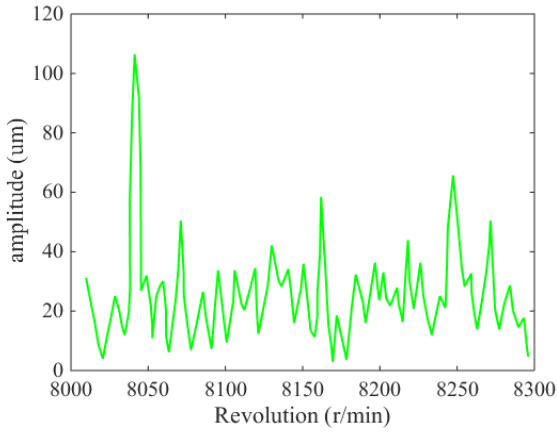
Fig. 6 CFF result of 27th blade at the rotational speed of 8043.2 r /min

Detailed results implying the vibration information such as resonant speeds, frequencies, and amplitudes of all the blades are shown in Fig. 7. It is found that all the 50 turbine blades resonate between the speed of 8043.2r/min and 8729.4 r/min and the resonance frequencies fluctuate between 1608.6 Hz and 1745.9 Hz. The maximum resonant amplitude is 118.5um. It also has a good match with the result by FE model analysis at 1st bending mode excited by 12EO.

Fig. 7 Detail vibration parameters of all the turbine blades on the same stage at 12EO resonance: (a) Amplitude; (b) Revolution; (c) Frequency.

B. 1ST Bending Resonance of 11EO

The results of the non-contact measurement also show that all the turbine blades resonate at the speed of 9100r/min roughly. The SDOF and CFF methods are also performed for the data processing to identify the key vibration parameters. The SDOF result of the 15th blade can be seen in Fig. 8 and the CFF result is shown in Fig. 9. Detailed results of resonant speeds, frequencies, and amplitudes of all the blades are shown in Fig. 10. All the results demonstrate that all the 50 turbine blades resonate in the speed range from 8953.4r/min to 9137.7r/min. The vibration frequencies vary from 1641.5Hz to 1675.2Hz. The maximum resonant amplitude is 25.1um. The 11 EO excites the 1st bending mode resonance, which is highly consistent with the results calculated by blade FE model in ANSYS.

In general, during the high-rotation and high-temperature operation, the centrifugal force increases the stiffness of the original blade and has a positive effect on the natural frequency while the accumulated thermal stress has the opposite effect. In the test, the above results from the non-contact measurement show that the natural frequencies of the turbine blades decrease as the rotational speed increases. It demonstrates that the thermal stress has a greater influence on the stiffness of the turbine blades than the centrifugal force. Therefore, the natural frequencies of the turbine blades reduce as the stiffness decreases resulting by the accumulating thermal stress when blades rotate.

For the vibration at the 11EO resonance, the amplitude stays on a relative low level, which affects little on the measurement accuracy with respect to the signal to noise ratio (SNR). However, for blade health monitoring and fatigue life evaluation, the blade resonance with small amplitude has trivial influence on blade operation. The resonance with considerable defection occurring in the operating speed greatly accounts on the severe failure of blades and should be taken great concern in the evaluation of blade high cycle fatigue.
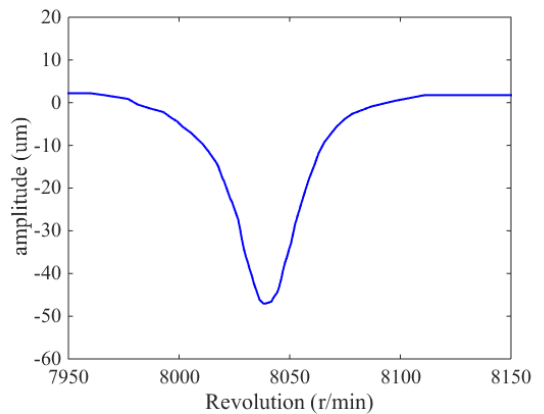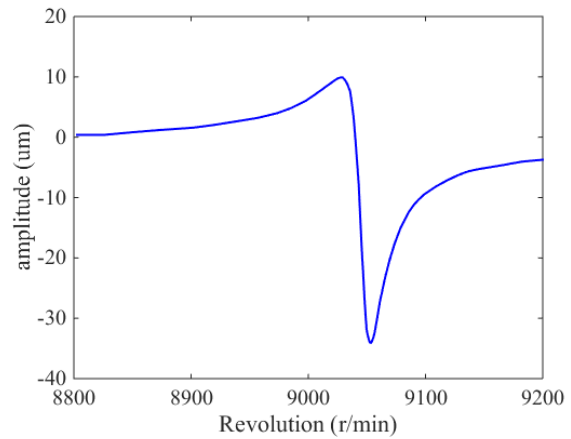

Fig. 8 SDOF result of 15th blade from 2th sensor at the rotational speed of 9048.9 r/min
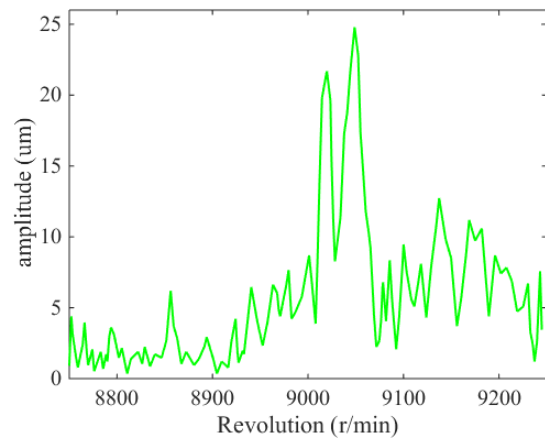
Fig. 9 CFF result of 27th blade at the rotational speed of 9100 r/min
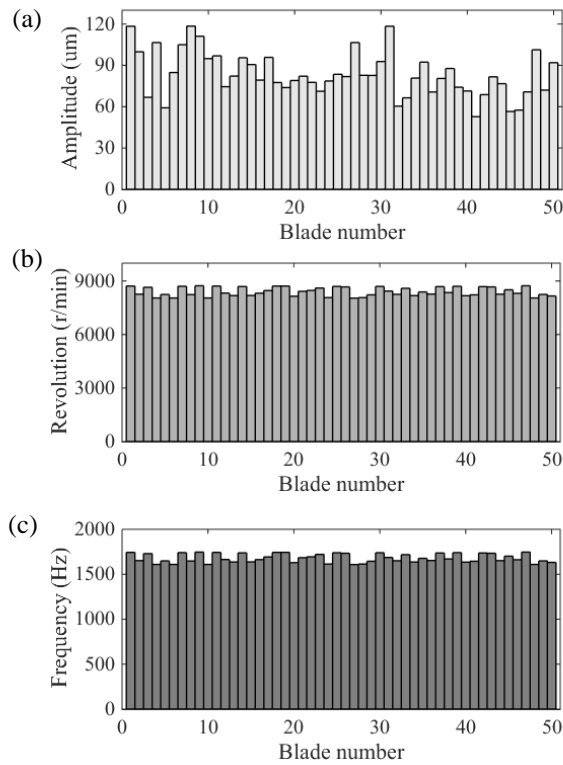
549
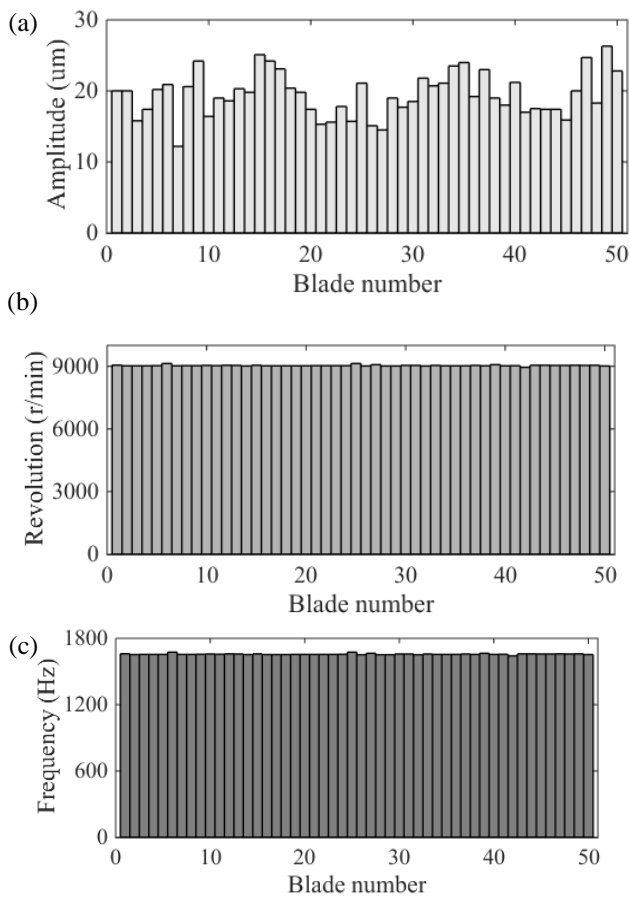
Fig. 10 Detail vibration parameter of all the turbine blades on the same stage at 11EO resonance: (a) Amplitude; (b) Revolution; (c) Frequency.

VI. CONCLUSION

In this paper, the optical sensors with cooling channels are designed and the air cooling methods are taken which lay the foundation of the tip-timing signal achievement. Then the proposed method for getting the once-per-revolution is innovative and simple which plays an important part for the data analysis. This method can be widely used in the non-contact vibration measurement of the aero-engine. The results of this paper prove that the non-contact vibration system based on tip-timing can be used in the turbine blades. The natural frequencies measured by the non-contact vibration system are very consistent with those of the finite element analysis method. It testifies that the non-contact vibration system will play a key role for on-line vibration measurement and analysis of the turbine blade.

ACKNOWLEDGEMENT

REFERENCES

[1] M. Mishra, J. Saari, D. Galar, U. Leturiondo, Hybrid models for rotating machinery diagnosis and prognosis estimation of remaining useful life. Technical Report, Luleå University of Technology, 2014.

[2] L. Liao, F. Köttig, Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery lifeprediction, IEEE Trans. Reliab. 2014, pp.191-207.

[3] J.Z. Sikorska, M. Hodkiewicz, L. Ma, Prognostic modelling options for remaining useful life estimation by industry, Mech. Syst. Signal Process. 25 (5) (2011) 1803–1836.

[4] G. Rigosi, G. Battiato, T.M. Berruti, Synchronous vibration parameters identification by tip timing measurements, Mech. Res. Commun. 79 (2017) 7-14.

[5] B. Salhi, J. Lardies, M. Berthillier, Identification of modal parameters and aeroelastic coefficients in bladed disk assemblies, Mech. Syst. Signal Process. 23 (6) (2009) 1894-1908.

[6] P. Russhard, Development of a blade tip timing based engine health monitoring system. University of Manchester, 2010.

[7] D.H. Diamond, P.S. Heyns, A.J. Oberholster, A comparison between three blade tip timing algorithms for estimating synchronous turbomachine bladevibration, in: In 9th WCEAM Research Papers, Springer, 2015, pp. 215-225.

[8] R.G. Du Toit , D.H. Diamond, P.S. Heyns, A stochastic hybrid blade tip timing approach for the identification and classification of turbomachine blade damage, Mech. Syst. Signal Process. 121 (2019) 389-411.

[9] M Mohamed, B P onello, P Russhard, A novel method for the determination of the change in blade tip timing probe sensing position due to steady movements. Mech. Syst. Signal Process. 126 (2019) 686-710.

[10] T Ouyang, W Guo, F Duan, New method for identifying rotating blades synchronous vibration based on tip-timing, Journal of Vibration and Shock, 30 (2011) 08. (in Chinese)

[11] M Liu, J Zhu, G Teng, Vibration measurement on compressor rotor blades of aero-engine based on tip- timing, Journal of Aerospace Power, 34 (9) (2019) 1895-1904. (in Chinese)

[12] M Liu, J Zhu, G Teng, The application of Non-contact vibration measurement system on the rotor blade vibration measurement of aero-engine. Gas Turbine Experiment and Research, 28(2)(2015) 45-48. (in Chinese)

[13] Z Zhong, G Teng, X Chen, G Li, Non-Contact vibration measuring system used for fan rotor blades. Gas Turbine Experiment and Research, 24(04)(2011) 44-47. (in Chinese)

[14] C Ao, B Qiao, M Liu, Y Sun, X Chen, Dynamic strain reconstructing method of rotating blades based on no-contact measurement. Journal of Aerospace Power, 35(3)(2020) 569-580.(in Chinese)

[15] K.K. Joung , S.C. Kang, K.S. Paeng , N.G. Park, H.J. Choi, Y.J. You, et al. Analysis of Vibration of the Turbine Blades Using Non-Intrusive Stress Measurement System, Asme Power Conference. 2006.

[16] K. Grant, Experimental Testing of Tip-Timing Methods Used for Blade Vibration Measurement in the Aero-Engine, Cranfield University, 2004.

[17] G. Dimitriadis, I.B. Carrington, J.R. Wright, J.E. Cooper, Wright, Blade-tip timing measurement of synchronous vibrations of rotating blade assemblies, Mech. Syst. Signal Process. 16(4) (2002) 599-62.

Distributed Space Remote Sensing and Multi-satellite Cooperative On-board Processing

Yang Liu*
Innovation Academy for Microsatellites
Chinese Academy of Sciences
Shanghai, China
lyhit1980@163.com

Yuanyuan Dai
Innovation Academy for Microsatellites
Chinese Academy of Sciences
Shanghai, China
mirrorelf@126.com

Guohua Liu
Innovation Academy for Microsatellites
Chinese Academy of Sciences
Shanghai, China
microsatlgh@163.com

Jingli Yang
Automatic Test and Control Institute
Harbin Institute of Technology
Harbin, China
jinglidg@hit.edu.cn

Longfei Tian
Innovation Academy for Microsatellites
Chinese Academy of Sciences
Shanghai, China
tianlf44@163.com

Hua Li
Innovation Academy for Microsatellites
Chinese Academy of Sciences
Shanghai, China
lh_secm@163.com

Abstract—**To meet the needs of emergency remote sensing applications, an intelligent remote sensing system based on distributed reconfigurable satellites was established, and research on multi-satellite on-board data fusion and intelligent processing technology based on distributed reconfigurable spacecraft was conducted. Taking the rapid detection and confirmation of the ship's on-board imaging target as an example, the multi-satellite on-board intelligent processing method is described in detail. Compared with the traditional remote sensing on-board data method, this method has the advantages of high recognition rate and low false alarm rate.**

Keywords—remote sensing, distributed, reconfigurable, on-board data fusion, on-board intelligent processing,

I. INTRODUCTION

With the development of aerospace technology and application requirements, the complexity and scale of individual spacecraft continue to rise, and it is inevitable to face launch difficulties, long development cycles, high costs, poor resistance to destruction, and certain specific space missions (such as global positioning and navigation, Ultra-long baseline synthetic aperture, etc.) cannot be completed by a single spacecraft, and the development of distributed spacecraft systems[1-3] becomes the future trend.

According to the single spacecraft size, distributed spacecraft can be roughly divided into two categories: module-level distribution and satellite-level distribution.

Typical studies of module-level distributed spacecraft include the F6 project, SBG (space-based group) and HISat cell stars. The F6 plan was proposed by the United States Defense Advanced Research Projects Agency (DARPA), and its purpose is to virtually form a fully-functional large satellite with the functional modules of separate flight. SBG was put forward in 2006, the basic concepts and methods are similar to F6. HISat Cellstar borrows biological and engineering principles to achieve a low-cost, modular, and scalable satellite platform structure that can support various types of payloads.

Satellite-level distributed spacecraft has a wide range of applications. In remote sensing, it can achieve a wide field of view, stereo detection, construct a virtual spacecraft to increase the focal length or expand the effective aperture (synthetic aperture) and other targets. Typical applications include A-Train (joint Detection), CANYVAL-X (using 2 cubic stars to form a virtual telescope to observe the corona, similar to Simbol x, Max, etc.), TanDEM-X and TechSat-21 (synthetic aperture radar), etc.

II. SATELLITE CONFIGURATION

The distributed reconfigurable space remote sensing system consists of 7 satellites, including a service star and 6 unit stars, which can serve the traditional remote sensing data application field and meet the needs of new applications such as emergency remote sensing, commercial remote sensing, and remote sensing big data mining. The research of the project will promote the high-precision, intelligence, networking and large-scale development of the next generation of microsatellites.

The system adopts an open satellite platform, which can achieve rapid satellite development and testing by building blocks. The satellite passes standard energy and mechanical locking interfaces and will meet different functional modules required by the mission, including power supply and power management, attitude measurement and modules such as control, satellite calculation and control, satellite-to-satellite data transmission and communication, inter-satellite data transmission and communication, structure and thermal control are assembled together. In order to ensure that the remaining remote sensing unit stars still have the ability to work collaboratively when there is a problem with the service star, each satellite has task planning and scheduling capabilities.
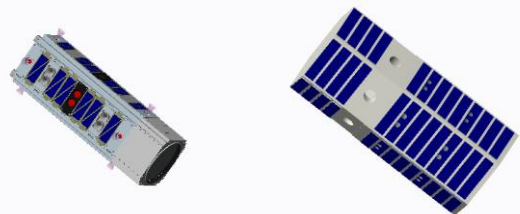

Fig. 1. Unit star and service star

The main characteristics of the satellite system are as follows:

- Satellite weight: no more than 15kg

- Attitude measurement and control

- Accuracy of three-axis attitude measurement: <10 ″

- Angular velocity measurement accuracy: <0.001 ° / s

551

- Angle control accuracy: <0.02 °

• Intelligent payload module

- Ground resolution: better than 4m (@ 500km)

- GNSS orbit positioning accuracy: better than 10m

- Width:> 16km (@ 500km)

- Image positioning accuracy without control point: better than 100m

• Multi-star system capabilities

- Long-term continuous monitoring mode: continuous monitoring duration> 12min

- Multi-star stitching imaging mode: stitching width 100km (@ 500km)

The satellite is modularly designed. The satellite prototype includes modules: intelligent payload module, structure, thermal control, attitude control module, integrated electronic module, route storage module, data communication sharing module, and distributed measurement and communication module. Different functional modules can be selected according to different needs.

A. High-precision Intelligent Payload Module

The high-precision intelligent payload module is composed of a visible light imaging camera, an attitude sensor, and a GNSS module. The visible light imaging camera is composed of an optical subsystem, a structural subsystem, and an electronics subsystem. It realizes the fusion of attitude determination, positioning data and payload data, and real-time control of autonomous imaging. It has the functions of ground-based visible area array imaging and video imaging.

B. Structure

The structure includes various structural parts of satellite and separation mechanism.

• Structural parts

It is used to bear and transmit the whole star payload, provide support, stiffness and dimensional stability for the satellite and related subsystems, provide a mechanical interface for the installation of on-board instruments and equipment, and provide protection for the on-board instruments and equipment.

• Separation mechanism

It is used to support satellite, provide mechanical connection interfaces for satellites and launch vehicles, and provide mechanical interfaces for satellite ground parking and transshipment.

C. Thermal Control Module

The inside of the star adopts isothermal design. The satellite bread is applied in multiple layers to form a heat insulation barrier inside and outside the star, reducing the influence of external heat flow on the temperature inside the satellite.

The use of multi-layer thermal insulation components and glass fiber reinforced plastic insulation gaskets between the satellite platform and the load module uses thermal control measures, so that the satellite platform and the load use independent heat dissipation channels and heat dissipation surfaces for independent temperature control to prevent mutual thermal effects.

D. Attitude Control Module

The attitude control module is composed of control components and actuators.

The task of the attitude control module is to achieve satellite attitude control, to meet the requirements of pointing accuracy and stability during the satellite mission: to eliminate the attitude disturbance deviation caused by the separation of the satellites and launchers after entering the orbit, complete the initial attitude determination, and achieve the basic sun pointing; complete the satellite control , orient the solar panel to the sun to ensure the energy supply. During the normal orbit of the satellite, overcome the various interference torques received by the satellite, ensure that the satellite reaches the pointing accuracy and attitude stability index requirements, and meet the imaging attitude requirements. According to the task requirements, implement attitude maneuver control to complete the switch between different directions.

The hardware part of the attitude control module is mainly composed of attitude sensor, attitude controller and actuator. Attitude sensor includes 1 set of fiber optic gyro component, 1 set of three-axis magnetometer, 1 set of sun sensor (including 6 sun sensor probes) and 2 star sensors, mainly used to measure a certain reference coordinate of satellite relative space department of posture. The attitude controller is used in the on-board computer to synthesize the measurement information of the sensor and solve the control quantity input of the actuator according to the control law. The actuator includes 3 reaction momentum wheels, 3 magnetic torque devices and 1 thruster.

E. Integrated Electronic Module (Including on Orbit Intelligent Processing, Computer, Power Controller):

It realizes the intelligent processing of data in orbit, the primary data processing function of unit satellite, the information extraction strategy of service satellite for different application scenarios, and the multi payload data processing and multi-dimensional information fusion.

The computer module realizes the satellite management, completes the whole satellite control, information processing, telemetry and remote control data receiving and sending. The fourth generation 65nm process of Actel is adopted. Smartfusion2 series on-chip system products based on flash architecture are used as processors, and ram, flash and other memory are expanded to meet the use requirements.

The power supply module provides the required electrical power for the equipment on the satellite in all operation stages and various working conditions during the life of the satellite to ensure that the power supply performance meets the requirements. According to the functional requirements, the distributed satellite power module is divided into the following functional modules: power management control and distribution unit (including 1 PCDU board), battery unit (including 1 battery board), and solar array.

F. Routing Module

Route and mass storage module (including hardware of inter satellite route exchange module) receive, store and forward payload data, satellite platform engineering telemetry data, inter satellite link data and other application data; complete coding, framing and distribution according to

protocol requirements. Realize fast networking route and service data exchange and sharing between satellites. According to the design idea of network information processing platform integration, the hardware design considers from the aspects of saving power and reducing volume, and explores how to make the platform complete multi-star information interconnection and data sharing. The hardware scheme adopts ARM + FPGA platform architecture which is widely used at present.

G. Data Communication and Sharing Module

Realize the data transmission and measurement and control functions between the satellite and the earth. The module fully inherits the existing product modularization and digital design ideas. It has the characteristics of low power consumption, small volume, light weight and perfect function. Data communication and sharing module mainly consists of antenna, RF transceiver channel, digital baseband and power supply.

The satellite can complete scan and video imaging functions separately. Multiple satellites combine to complete wide splicing, continuous tracking and stereo imaging mode.

Fig. 2. wide splicing mode

Fig. 3. continuous tracking mode

Fig. 4. stereo imaging mode

III. On-Board Data Fusion And Intelligent Processing

The automatic detection of ship targets on the sea surface is widely studied in the field of remote sensing application. At present, there has been a successful on orbit processing system. However, at present, the high detection rate (such as more than 90%) is achieved at the expense of the false alarm rate index. In panchromatic images, broken clouds, waves, etc. are very easy to be detected as false alarms, so that the actual detected targets are often many false targets, which brings great trouble to the follow-up information utilization. From this point of view, the detection and confirmation of ship target is still a key issue.

The distributed reconfigurable spacecraft adopts the aggregate observation mode, that is, the seven distributed satellites that make up the spacecraft adopt the aggregate formation mode to realize the cooperative observation of the same area at the same time. In this paper, the working mode of the spacecraft is defined as the simultaneous cooperative observation mode. For the application of sea target tracking and motion situation awareness, the distributed reconfigurable spacecraft adopts the aggregate formation mode Sequential observation method is adopted, that is to say, seven distributed satellites of the spacecraft are in series formation mode, and the satellites are separated by a certain distance, and the time-sharing cooperative observation of the same area is realized by successively passing over the same observation area.

The project mainly solves the problems from the following aspects: first, make full use of the spacecraft's ability of distributed multiple observation, effectively extract the space domain image information and time domain motion information of ship targets, and remove the false alarms such as broken clouds and broken waves that do not meet the constraints of space and time information. Specifically, in the distributed orbit processing and fusion computing framework, the unit star node can carry out the traditional idea of "replacing high false alarm with high detection rate" to ensure that the real ship targets are not missed and suspected ship targets are detected as much as possible, while the service star node can remove the isolated false alarm targets and the "exposed" false alarm targets in the space-time domain to improve the detection accuracy through fast association and fusion identification of these suspected ships, and previous research results have shown that through fusion it can greatly improve the confidence level of detected targets. Secondly, the ship target detection method currently applied in orbit only filters for large clouds, but less distinguishes between the details of ships, broken clouds and broken waves, strengthens the organization and utilization of the details, and reduces the ship target from a single data source by establishing more complex ship target identification features and the false alarm rate level of target detection. Thirdly, through the satellite ground collaborative processing service mechanism, it provides a priority for the detection of ship targets, and through the offline learning and on orbit updating and reconstruction on the ground, it is also a very effective technical way to improve the accuracy level of on orbit processing and fusion module.

The technical process is shown in the figure below. It can be seen that the intelligent processing links distributed on the unit satellite include: cloud judgement [4,5], sea land separation [6,7], rapid detection of suspected ship target based on the target [8,9], target classification identification and target location solution. The distributed intelligent processing links on the service star include: target association, ship target determination, etc. These links in series constitute the technical framework of rapid detection and fusion of sea targets [10-13]. The subject needs to study the algorithm of these core links.
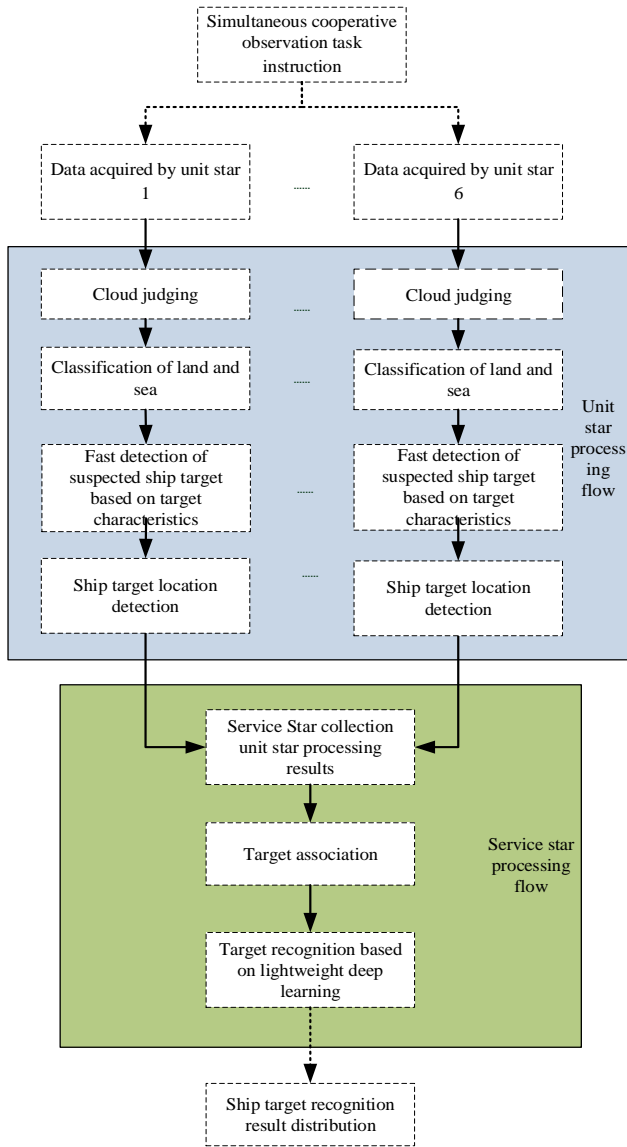
Fig. 5.Technical framework of fast detection and fusion of sea targets

A. Cloud Judging Technology

Cloud discrimination is one of the most important problems in remote sensing data processing. Because of the existence of cloud cover, the utilization rate of remote sensing data is reduced, so it is difficult to ensure the accuracy of target recognition and classification, sometimes even impossible to carry out, so it is very important to carry out cloud discrimination. The framework of feature extraction and classifier recognition is proposed for cloud identification, which requires not only accurate location of cloud image area, but also low complexity of algorithm and real-time performance.

The application background of cloud judgment technology puts forward many special requirements for its feature extraction and selection

- Due to the diversity of cloud layer and the complexity of surface features, it is often necessary to use a variety of features for joint discrimination to increase the separability of cloud and other surface features in the feature space.

- Cloud judgment often needs to sample remote sensing data in order to reduce the amount of calculation,

which requires that the features will not change due to the increase or decrease of image scale and meet the scale invariance.

- Due to the real-time requirements, the complexity of feature calculation is required to be low.

In pattern recognition system, the most commonly used features include both spatial and frequency domain features. Spatial domain features include mean value, histogram sum, histogram variance, texture co-occurrence matrix, histogram equalization, etc., while frequency domain features include cosine transform, fourier transform, wavelet transform, etc. Considering the effectiveness and real-time, the following four features are considered:

Image mean value

$$\bar{\rho} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} f(i,j) \tag{1}$$

Image variance

$$\sigma = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(f(i,j) - \bar{\rho} \right)^2 \tag{2}$$

Image edge

$$S_d = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(|f(i,j) - f(i+1,j)| + |f(i,j) - f(i,j+1)| \right) \tag{3}$$

Image entropy

$$\xi_e = -\sum_{m=0}^{255} P(m) \log P(m) \tag{4}$$

Where p (m) represents the probability that a pixel with gray level equal to m appears in the image block. These four features are calculated based on image blocks of the same scale. When the remote sensing data arrives, it is first sampled and divided into equal size image blocks, and then the four features are calculated as the basic decision unit to form the feature vector

$$v = [\bar{\rho} \quad \sigma \quad S_d \quad \xi_e]$$

Then the feature vector is mapped to a point in the feature space as the basis of cloud image discrimination. The specific process of cloud judgment is as follows:

Step1: block the acquired image to form an image block of equal size (n * n) as the basic decision unit.

Step 2: four features are extracted from each image block unit to form a four-dimensional feature vector.

Step 3: in the feature space, calculate the Euclidean distance between the feature vector and cloud ground clustering center.

Step 4: take the classification of the minimum distance as the judgment result of whether the basic unit is cloud or ground object.

Step 5: conduct connectivity analysis to remove cavities and outliers.

B. Separation Technology of Land and Sea

Due to the difference between the features and algorithms of sea target detection and land target detection, sea land separation is of great significance for the subsequent target

detection and recognition. Sea land separation can be realized by means of the longitude and latitude geographic information stored in the satellite, and the longitude and latitude of the pixel can be used to determine whether the corresponding pixel is land or sea area, so as to realize the sea land separation; it can also be realized by image processing algorithm, generally speaking, because the accuracy of the extrapolated longitude and latitude information on the satellite is not high, so it needs to use two methods together.

The technology consists of three processes. Firstly, the coastline information provided by the marine geographic information database is used to realize the preliminary separation of the sea and land areas. Because the resolution of the marine geographic information database is low, it is necessary to combine the image features of different areas of the sea and land, use the image of the sea, and extract the multi-channel image features of the sea and land based on the sample learning method At last, the output of cloud judgment is integrated to provide reliable and effective template information for subsequent processing.

C. Fast Detection Technology of Ship Target

In view of the obvious difference between the bottom visual features of the ship and the sea background, the method of significance detection is used to determine the candidate target area. Because the ship targets on the sea have good closed contour, the module adopts a fast extraction method based on the target. Firstly, according to the scale of different ship targets, the image is zoomed to multiple different scales (generally 3); then, the method of convolution calculation is used to extract the target to speed up the detection speed; after convolution linear mapping, the module can get the discriminative characteristics of the target area and the non-target area, the classifier of the target and the non-target can determine whether the window area has the "target"; through the non-maximum suppression, the multiple repeated targets of the same area can be eliminated; finally, the location and size of the "target" area can be obtained by reflecting the "target" probability map on the original image to realize the rapid detection of the suspected ship target area.

D. The Technology of Ship Target Location

The essence of ship positioning is to calculate the longitude and latitude of the target on the image according to the camera imaging parameters, camera attitude parameters, satellite attitude parameters, satellite orbit parameters and the earth elevation information.

In satellite photogrammetry, the space position and ground point coordinate of satellite are usually expressed in the geocentric rectangular coordinate system, while the attitude angle of sensor is defined as the rotation angle of satellite body coordinate relative to its orbit plane. Due to the conversion of multiple spatial coordinate systems, it is very complex to establish a strict geometric processing model of satellite remote sensing image. It is inevitable to use coordinate conversion between multiple spatial coordinate systems, and the definition of conversion parameters and their mathematical forms will change with the different types of sensors.

By defining the ground photogrammetry coordinate system parallel to the image space auxiliary coordinate system, the collinear condition equation can be directly expressed as the mathematical relationship between the image space coordinate of the image point and the photogrammetry

coordinate of the corresponding ground point. The external orientation of the image is based on the photogrammetric coordinates. The image longitude and latitude calculation module calculates the geographic coordinates of the image point corresponding to the ground point according to the mathematical relationship between the spatial coordinates of the image point and the photogrammetric coordinates of the corresponding ground point.

The calculation of longitude and latitude involves a series of coordinate system conversion processes: camera coordinate system - > satellite body coordinate system - > satellite orbit coordinate system - > geocentric rectangular coordinate system image coordinate system. Firstly, the imaging equation from image point to ground is established, and then the ellipsoid equation and elevation information of the ground are combined to solve the longitude and latitude of the ground target point corresponding to image point through iterative calculation.

E. Ship target fusion Confirmation Technology

Target fusion confirmation is to make a decision on the real existence of the target through the ship target detection results and information association results of multiple unit satellite sources, especially in the process of accurate situation awareness in the sea battlefield, it can reduce the false alarm and improve the confidence level of the detection results, which is of great significance.

When the target fusion of multiple sources is confirmed, if all sources can be ensured, the framework of data fusion recognition between them is generally as shown in the following figure:
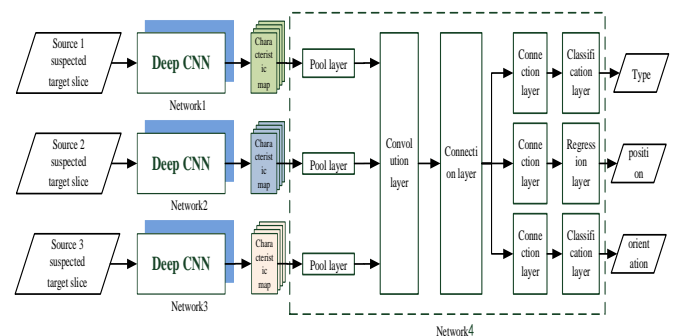
Fig. 6. Target recognition technology framework of multi-source image fusion based on deep learning

The slice images of suspected targets provided by each source are identified based on the lightweight depth recognition neural network. The results are fused according to the scores of corresponding modes to get the final recognition results. Because of the visible light load of each unit star and the same imaging parameters such as orbit height and resolution, the lightweight depth neural network target recognition can be carried out in parallel on the service star. After processing, we can get the identification results of each source data, and then use the mode fusion method to achieve the final confirmation.

IV. CONCLUSION

the intelligent processing technology of on orbit data has very clear purpose and important research significance. The distributed orbit processing system takes the multi-satellite collaboration as the core, and transfers independent imaging information through the inter satellite links. The method is

advanced and of great application value. It can provide the technical foundation for the future intelligent network application.

REFERENCES

[1] Tapan Sharma, Vinod Shokeen, Sunil Mathur. Distributed Approach to Process Satellite Image Edge Detection on Hadoop Using Artificial Bee Colony. 2020, 11(2):80-94.

[2] Chong Wang,Jinhui Tang,Xiaohang Cheng,Yingchen Liu,Changchun Wang.Distributed cooperative task planning algorithm for multiple satellites in delayed communication environment[J].Journal of Systems Engineering and Electronics,2016,27(03):619-633.

[3] Wongoo Lee,Yunsoo Choi,Kangryul Shon,Jaesoo Kim.Fast distributed and parallel pre-processing on massive satellite data using grid computing[J].Journal of Central South University,2014,21(10):3850-3855.

[4] Proia, N. and V. Pagé, Characterization of a bayesian ship detection method in optical satellite images[J]. Geoscience and Remote Sensing Letters, IEEE, 2010. 7(2): p. 226-230.

[5] Corbane, C., et al., A complete processing chain for ship detection using optical satellite imagery[J]. International Journal of Remote Sensing, 2010. 31(22): p. 5837-5854.

[6] Liu, G., et al., A New Method on Inshore Ship Detection in High-Resolution Satellite Images Using Shape and Context Information[J]. Geoscience & Remote Sensing Letters IEEE, 2014. 11(3): p. 617-621.

[7] Yang, G., et al., Ship Detection From Optical Satellite Images Based on Sea Surface Analysis[J]. IEEE Geoscience & Remote Sensing Letters, 2014. 11(3): p. 641-645.

[8] Shi, Z., et al., Ship Detection in High-Resolution Optical Imagery Based on Anomaly Detector and Local Shape Feature[J]. IEEE Transactions on Geoscience & Remote Sensing, 2014. 52(8): p. 4511-4523.

[9] Lin, J., et al., A line segment based inshore ship detection method, in Future Control and Automation. 2012, Springer. p. 261-269.

[10] Jing L H, Cheng Q M. An Image Fusion Method Based on Object-oriented Classification[J]. International Journal of Remote Sensing, 2012, 33(8): 2434-2450.

[11] Meng Q M, Borders B, Madden M. High-Resolution Satellite Image Fusion Using Regression Kriging[J]. International Journal of Remote Sensing, 2010, 31(7): 1857-1876.

[12] Ekhtari N, Zoej M J V, Sahebi M R. et al. Automatic Building Extraction from LIDAR Digital Elevation Models and World View Imagery[J]. Journal of Applied Remote Sensing, 2009, 3(1): 033571.

[13] Jin X Y, Davis C H. An Integrated System for Automatic Road Mapping from High Resolution Multi-spectral Satellite Imagery by Iinformation Fusion[J].Information Fusion, 2005, 6(4): 257-273.

A Method of Antenna Gain Testing without Standard Gain Antenna

Kuo Gao*
Shanghai Engineering Center for Microsatellites
Innovation Academy for Microsatellites, Chinese Academy of Sciences
Shanghai, China
gaok61096@163.com

HuiJie Liu
Shanghai Engineering Center for Microsatellites
Innovation Academy for Microsatellites, Chinese Academy of Sciences
Shanghai, China
hjliu72@hit.edu.cn

Lei Liu
Shanghai Engineering Center for Microsatellites
Innovation Academy for Microsatellites, Chinese Academy of Sciences
Shanghai, China
liul@microsat.com

XiaoTong Pan
Shanghai Engineering Center for Microsatellites
Innovation Academy for Microsatellites, Chinese Academy of Sciences
Shanghai, China
panxt@microsat.com

Hao Wang
Shanghai Engineering Center for Microsatellites
Innovation Academy for Microsatellites, Chinese Academy of Sciences
Shanghai, China
wangh@microsat.com

Yang Liu
Shanghai Engineering Center for Microsatellites
Innovation Academy for Microsatellites, Chinese Academy of Sciences
Shanghai, China
lyhit1980@163.com

Abstract—The gain calibration value of the standard gain antenna is the reference baseline of the unknown gain antenna measurement, which determines the accuracy of the measured antenna gain measurement. This paper introduces a method of measuring antenna gain without standard gain antenna. Taking horn antenna as an example, a set of antenna measurement system with unknown gain is established in microwave darkroom by using three pairs of horn antenna, vector network analyzer, tripod and other equipment. Taking one antenna as a reference, the gain values of the other two antennas are measured respectively, and the actual measured values are compared with the gain values in the antenna manual. The gain deviation is within a reasonable range, which shows that the method is reasonable and feasible when there is no standard gain antenna.

Keywords: lack of, standard gain antenna, horn antenna, gain measurement

I. INTRODUCTION

Horn antennas are widely used in the field of microwave testing, and are often used as standard gain antennas to measure the index of measured antenna. The accuracy of the measured antenna gain index is determined by the accuracy of the standard gain value of the standard gain horn antenna [1]. However, at present, it is difficult for all domestic measurement agencies to ensure accurate calibration of standard gain antennas. If a standard gain antenna with calibration data from a measurement agency is imported from abroad, the price of the calibration data at a single frequency point is even higher than the antenna under test itself [2], very expensive, so at present most domestically use the theoretically calculated gain value of the standard gain horn antenna as its standard gain value. Although the horn antenna gain measurement method given in document[3] does not use the standard gain horn antenna, it has an ideal assumption when measuring the horn antenna gain, the two horn antenna gains are exactly the same, which is It is impossible to have two antennas with exactly the same gain index.

The author often uses the horn antenna to perform related RF signal tests. In many cases, it is necessary to use the accurate gain value of the horn antenna at the specific frequency used. At this time, it is necessary to calibrate the gain value of the used horn antenna. And the author's unit does not have a standard gain antenna with calibration data of the measurement mechanism, so a method for measuring the gain of the horn antenna without a standard gain antenna is proposed, which are used to measure the gain index of the antenna with unknown gain.

II. ANTENNA GAIN MEASUREMENT SYSTEM

A. Theoretical Support for Antenna Gain Measurement

For most antenna users, what is usually required is the antenna's far-field radiation characteristics. The antenna gain (without special instructions) generally refers to the antenna's far-field radiation gain. The far-field gain test for antennas theoretically requires infinite test distance, which is unrealistic in real life. For accurate measurement and operability of antenna gain, the engineering implementation process usually selects the appropriate test distance, which is generally recognized and the minimum test distance in the far field that is actually applied is [4-6] formula (1).

$$R_{min} = 2D^2 / \lambda \qquad (1)$$

In formula (1), R represents the distance between the surface of two antennas, unit is meter; D is the maximum size of the antenna under test, which is the sum of the maximum size of the antenna under test and the maximum size of the auxiliary antenna, unit is meter; λ is the working wavelength of the antenna, usually the largest antenna operating wavelength is selected as the value, unit is meter.

The operating frequency of antenna is 1~3GHz, So λ selects the wavelength when the frequency is 3GHz; the three horn antennas which are selected have the same size, the maximum size is d = 0.2712 meters, and the minimum test distance of this experiment can be obtained:

$$R_{min} = 2 \times (d+d)^2 / (c/f) = 2 \times 0.5423^2 \times 3 \times 10^9 / 3 \times 10^8$$
$$= 5.889m \qquad (2)$$

According to the microwave signal power transmission formula in document [4,5], the actual measurement of the horn antenna gain is based on this formula.

$$P_r = (\lambda / (4\pi D))^2 P_t G_t G_r \qquad (3)$$

Where: P_r is the maximum receiving power of the receiving antenna;

P_t Input power of the transmitting antenna;

G_t Is the gain of the transmitting antenna;

G_r Is the gain of the receiving antenna;

D is the distance between the receiving and transmitting antenna ports;

λ Is the signal wavelength.

Take the logarithm to base 10 on both sides of formula (3), multiply by 10 to get:

$$10\log P_r =$$
$$20\log\left(\lambda/(4\pi D)\right)+10\log P_t + G_t' + G_r' \qquad (4)$$

Move items merged:

$$10\log P_r - 10\log P_t =$$
$$20\log\left(\lambda/(4\pi D)\right)+G_t'+G_r' =$$
$$20\log\left(c/(4\pi fD)\right)+G_t'+G_r' =$$
$$20\log\left(c/(4\pi)\right)-20\log f -20\log D + G_t'+G_r' =$$
$$147.558-20\log f -20\log D + G_t'+G_r'$$
$$(5)$$

f is the signal frequency, the unit is Hz;

D is the distance between the transmitting and receiving antenna ports, the unit is meter;

G_t' is the gain of transmitting antenna, the unit is dBi;

G_t' is the gain of receiving antenna, the unit is dBi.

Assume that:

$$L_f = 147.558-20\log f -20\log D =$$
$$138.0156-20\log f$$

L_f represents the free space loss value of different frequency. Through calculation and simulation, the FSL (free space loss) surface of microwave signal with frequency and distance is plotted as shown in Figure 1. From the simulation results, it can be seen that the FSL value of the microwave signal increases with the increase of the spatial distance, and increases with the increase of the signal frequency.
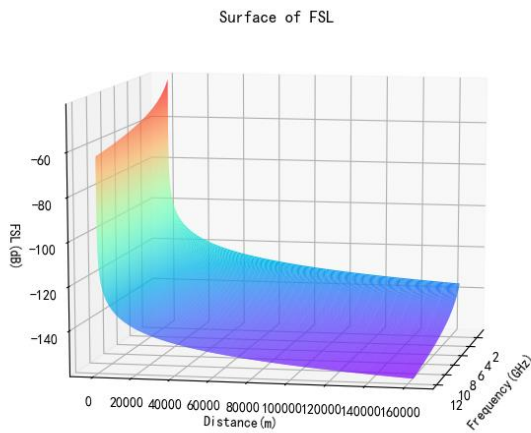
Fig.1. Space loss surface of microwave signal with different frequency varying with distance

Considering comprehensively the accuracy of the experiment, reducing the system error, and the physical size

of the microwave dark room, this measurement method measures the horizontal distance D=6 meters and 7 meters of the two horn antennas respectively. The gain value in the horn antenna factory manual is compared to verify the effectiveness of this method; the horn antenna is fixed with a tripod and the height is H = 1.5 meters; according to the formula $L_f =147.558-20\log f -20\log D$, Let D = 6/7/8 to obtain the free space loss value that varies with frequency when the distance between the antenna surface is 6/7/8 meters. Calculate and draw the curve as shown in Figure 2:
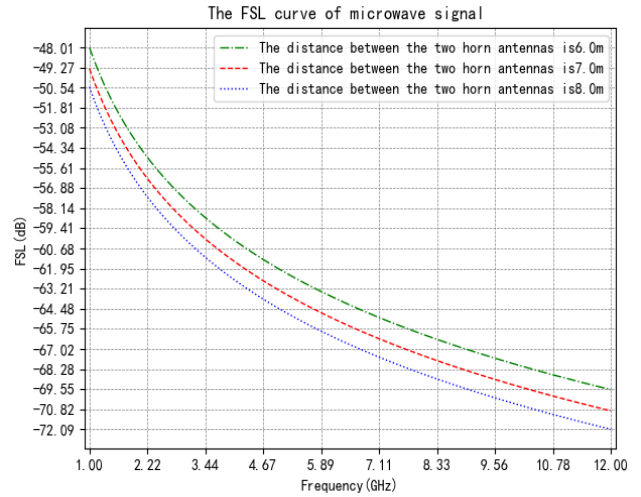
Fig. 2. The free space loss of microwave signal varies with frequency at distance of 6-8 meter

B. Three Horn Antennas Measurement System

The equipment and accessories used in the three horn antennas measurement system include a darkroom of microwave, Gain to be measured horn antenna A and B, a reference horn antenna C with unknown gain, two tripods, a vector network analyzer, and two 10 meters Long high-frequency cables, several RF adapters, etc. The connection diagram of the experimental scene is shown in Figure 3 and Figure 4.
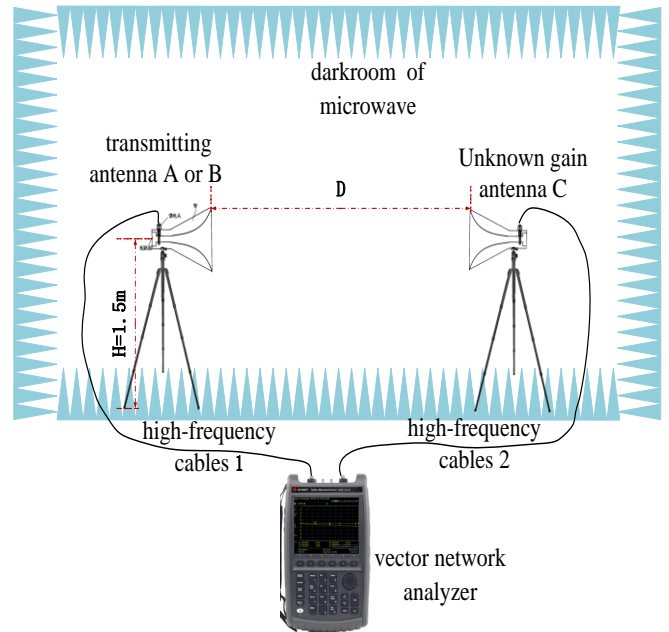
Fig.3. Connection block diagram of three horn antenna gain measurement scene (part 1)
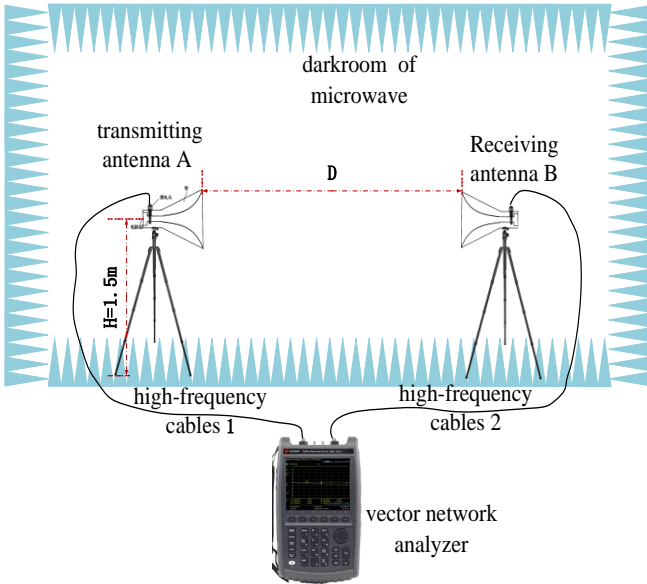
Fig. 4. Connection block diagram of three horn antenna gain measurement scene (part 2)

Figure 5 shows the actual measurement scenario of the three-horn antenna gain measurement system (Part 1). Two of the horn antennas used are dual-spine horn antennas with unknown gain. The other horn antenna has the antenna gain curve given in the antenna manual. The vector network analyzer model is keysight N9918A.
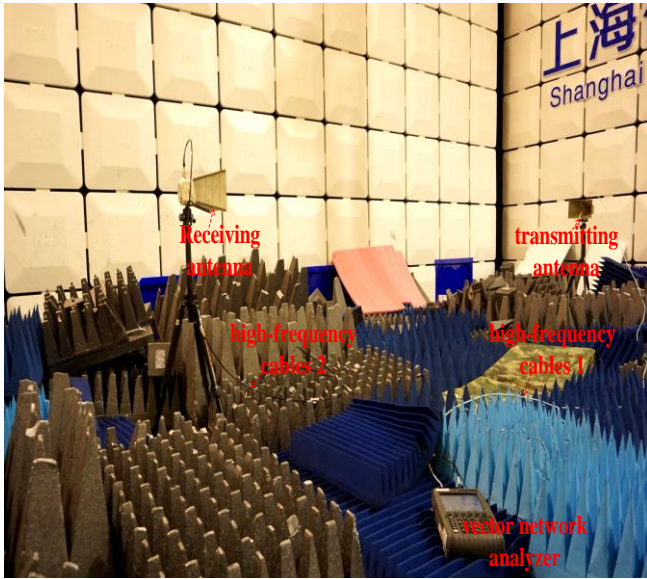
Fig. 5. The measurement scene of horn antenna gain measurement system(part 1)

III. METHOD OF HORN ANTENNA GAIN MEASUREMENT

A. Zero Calibration of The Measuring System

Before the actual test starts, Calibration of cable and vector network analyzer. The working frequency range of the horn antenna to be tested is 1~3GHz. So set the frequency range of the vector network analyzer to 1~3GHz, the high-frequency cable at the antenna port of the horn to be tested is directly connected with a microwave adapter, and the loss caused by the vector network itself and the high-frequency cable used is zeroed by the calibration function of the vector network analyzer Cal. The specific link method for zero calibration is shown in Figure 6.
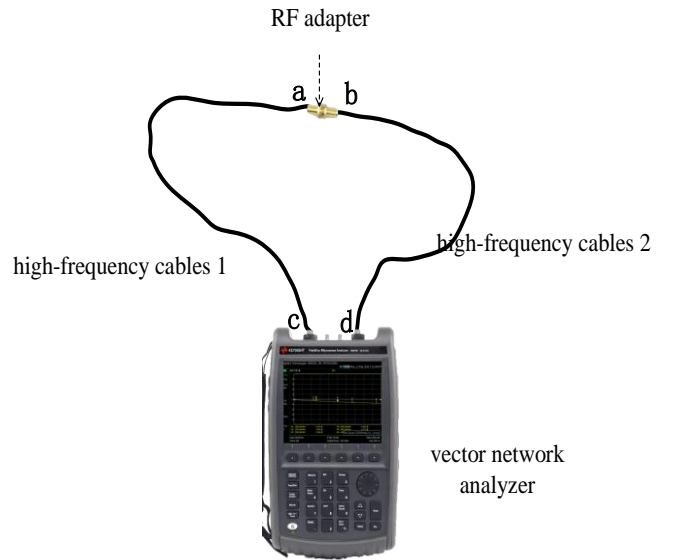
Fig. 6. Calibration connection diagram of vector network analyzer

B. Reference Gain Measurement

The two horn antennas with gain to be measured are referred to as antenna A and antenna B, and the reference horn antenna with unknown gain is referred to as antenna C. Antenna A and antenna B are in the same position, antenna A and antenna C, antenna B and antenna C, time-sharing test, establish an experimental environment according to the connection of scene in Figure 3, and measure the corresponding S21 curve, the two sets of real test results S21_AC and S21_BC are saved to CSV files respectively, which is convenient for post-processing to get the antenna gain, some results are shown in Table 1.

TABLE 1 Part of S21 measured results recorded by vector network analyzer

frequency(GHz)	S21_AC(dB)	S21_BC(dB)	S21_AB(dB)
1	-33.02443	-35.6444	-34.3829
1.055	-33.3575	-33.4398	-33.7632
1.11	-33.6956	-35.5616	-34.7755
1.165	-33.2114	-34.3335	-32.0278
1.22	-35.3633	-35.5552	-34.1547
⋮	⋮	⋮	⋮
2.87	-33.9472	-34.9597	-34.7028
2.925	-34.3943	-34.7166	-34.5664
2.98	-33.8312	-35.1506	-34.5621
3.035	-34.1228	-34.7456	-33.8041

According to formula (5) and vector network analyzer test results in Table I, Draw the following conclusions:

$$S21\_AC = L_{f-6,7} + G_C + G_A \qquad (6)$$

$$S21\_BC = L_{f-6,7} + G_C + G_B \qquad (7)$$

In the formula, S21_AC and S21_BC are the S21 values tested when the distance between the two antennas is 6 meters or 7 meters, $L_{f-6,7}$ is the FSL values at the corresponding frequency when the distance between the two antennas is 6 meters or 7 meters, G_A, G_B and G_C represent the gain values of antennas A, B and C, respectively.

formula (6) minus formula (7)

$$S21\_AC - S21\_BC = G_A - G_B \qquad (8)$$

C. Relative Gain Measurement

Establish the experimental environment according to the scene connection in Fig4 and replace the position of antenna C in Fig3 with antenna A. The position of antenna B constant antenna B in Fig3. The distance between antenna A and antenna B is 6 meters and 7 meters. The corresponding S21 curve is measured with a vector network analyzer, and the test result S21_AB is saved to a CSV file to obtain the relative antenna gain of antenna A and antenna B.

According to formula (5) and vector network analyzer test results in Table 1, Draw the following conclusions:

$$S21\_AB = L_{f-6,7} + G_B + G_A \qquad (9)$$

IV. RESULTS COMPARISON

According to the free space loss results at the distance of 6 meters and 7 meters in Figure 2 and formula (8) and formula (9), the absolute gain of the antenna A and antenna B to be tested can be obtained.

$$G_A = (S21\_AB + S21\_AC - S21\_BC - L_{f-6,7})/2 \qquad (10)$$

$$G_B = (S21\_AB - S21\_AC + S21\_BC - L_{f-6,7})/2 \qquad (11)$$

According to the gain curve in the manual of the horn antenna, combined with the test results at two different distances in this actual test, the final measured gain of the horn antenna to be tested is compared with the gain data value in the antenna manual, as shown in Figure 7.
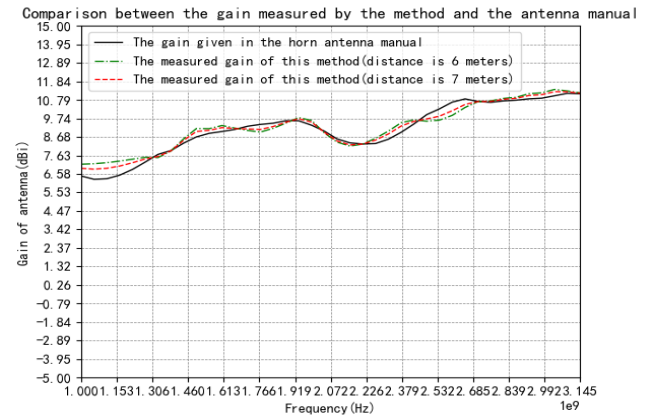
Fig. 7. Comparison between the measured gain of horn antenna and the gain in antenna manual

V. CONCLUSION

In the darkroom of microwave, in the absence of a standard gain horn antenna, use a vector network analyzer and three horn antennas with unknown gains to perform ingenious combination tests, and the method of using the computer as an auxiliary tool to measure the horn antenna gain is reasonable and feasible. Under the conditions of the antenna far-field test, two distance test scenarios are selected. The comparison between the actual test of the unknown gain and the gain value given in the antenna manual is within a small error range. This method can provide engineering reference for the measurement of horn antenna gain without a standard gain horn antenna.

REFERENCES

[1] Gary E. Ewans, Antenna Measurement Techniques, Artech House, Inc, PP. 209-225, 1990.

[2] WRIXON G T, WELCH W J. Gain measurement of standard electromagnetic horns in the K and Ka bands[J]. IEEE Transactions on Antennas and Propagation, 1972,20(3):136-142.

[3] Zhou Yandong, Liu Zhen, Liu Rubing, and Lin Qi, Horn Antenna Gain Measurement Based on Vector Network Analyzer, Science and Technology Innovation Herald, 2017 NO.03:92-95.

[4] Agilent AN 1287-2. "Exploring the Architectures of Network Analyzers". Agilent Application Note.

[5] L.A.Muth and R.L.Lewis, "A General Technique to Correct Probe Position Errors in Planar Near-Field Measurements to Arbitrary Accuracy". IEEE Transactions on Antennas and Propagation, Vol.38, No.12, PP.1925-1932, December 1990.

[6] Y.Rahmat-Samii et al. "The UCLA Bi-Polar Planar-Near-Field Antenna Measurement and Diagnostics Range". IEEE Antennas and Propagation Magazine, December 1995, Vol.37, No.6, pp.16-35.

Design and Application on Complicated Power System Operation Cockpit Technologies

Zhiyong Li *
Power Dispatching & Control Center
China Southern Grid
Guangzhou, China
lizhiyong@csg.cn

Chunxiao Liu
Power Dispatching & Control Center
China Southern Grid
Guangzhou, China

Abstract—**Power system operation cockpit(POC) is the integration of operation smart system(OS2). After 7 years of research and application, it has been popularized in CSG. At the same time, the original design concept of POC has been landed and become an important platform in power system operation. Both the intension and the extension of POC has been enriched by the test of actual application. This paper focuses the framework, function and practical benefits of POC after a profiles of its position in OS2. This paper is a system application summary of POC, which is aimed at supporting the development of power system automatization.**

Keywords—power system operation cockpit; operation smart system; system structure; function design; practical benefit.

I. INTRODUCTION

In recent years, with the rapid development of power system, the interconnection of large power grids makes the scale of power grids increasing and the complexity of operation increasing day by day, which puts forward higher requirements for dispatching and operation command. In 2011,in order to meet the requirements of power grid development, the concept of Operation Smart System (OS2) was put forward in Southern Power Grid, and it was popularized and applied in three-level dispatching agencies of network, province and prefecture according to the principle of "six unifications"[1-3].

Power System Operation Cockpit, as the core part of OS2 system, is the key framework to support dispatcher dispatching and operation command. In the framework of OS2 system, document [4] presents a preliminary framework to implement an Operation Cockpit, including data format, architecture and module design of the system.On this basis, document [5] investigate on the key performance indicators (KPI) for self-healing systems and operation cockpit based on balanced scorecard (BSC). Documents [6] presents a indices-based dispatching support system which provides a hierarchical multi-level multi-department structure to combines a management-level and technical level to the system operators. Document [7] introduces IBM Intelligent Operations Centre (IOC) which includes incident management, integrated mapping and collaboration capabilities.

As an important part of OS2, the current research on the operating cockpit is still focused on the functional design level. In fact, since the concept was put forward, cockpit operation in the South Network has been deeply studied and widely applied, solved a series of important problems in production and operation, and achieved remarkable

---

* LI Zhiyong is the corresponding author. (e-mail: lizhiyong@esg.cn).

application benefits.

According to the above ideas, this paper will further sort out the functional orientation of the operating cockpit in OS2 system architecture, introduce the frame design and system design of the operating cockpit in detail, and finally analyze its application benefits and summarize the mature experience of popularization and application combined with its application in South Network.

II. FUNCTIONAL POSITIONING OF OPERATING COCKPIT

A. Architecture Framework of OS2

OS2 system of South Network will cover the four-level dispatching agencies of network[8-9], provinces, prefectures and counties, and will extend to power plants, substations, patrol centers, monitoring centers and other regulatory objects to form a unified standard and complete dispatching business support system. The main station system in network, province and prefecture is the core of OS2 system, and its architecture is shown in Figure 1.
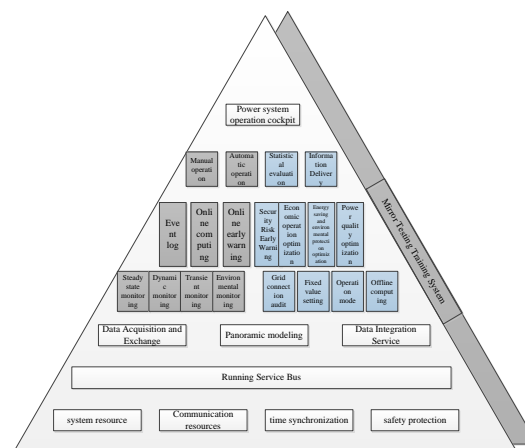
Figure1. Framework of master station system

The main station system can be divided into eight business modules in terms of function, which are operation cockpit, four application centers, operation service bus, basic platform and mirror test and training system. The four application centers are intelligent data center, intelligent monitoring center, intelligent control center and intelligent management center. Under the above system, the operating cockpit, as the top level of the main station system, shoulders the responsibility of information exchange between OS2 system and dispatching operators.

In fact, OS2 master station system can further integrate the four application centers into Operation Control System

(OCS), Operation Management System (OMS) according to their relationship with operation control and production management. The main station architecture shown in Figure 1 can be simplified as shown in Figure 2. The Intelligent Monitoring Center and Intelligent Control Center in the four application centers are integrated into the operation control system, and the Intelligent Management Center is integrated into the operation management system. The dispatcher interacts with the cockpit to grasp the operation information of the power grid and issue the dispatching operation control requirements to realize the dispatching operation management.
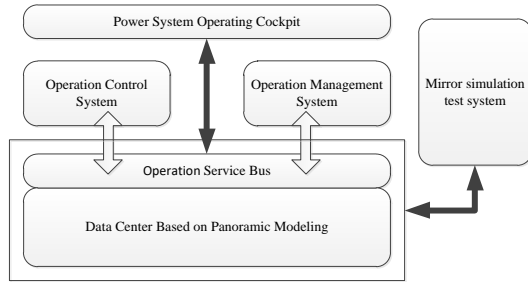
Figure 2. Location of POC

B. Dsesign Concepts and Solutions of Power System Operating Cockpit

Cockpit is the center of operation control system, which provides a platform for centralized monitoring and operation for pilots. It has been widely used in spacecraft, aircraft and other operation control systems.

Introducing the concept of cockpit into power system operation control, its starting point is to construct a centralized operation control platform, to order the huge information of power grid operation, to provide scientific and accurate decision-making basis for dispatchers and operators based on in-depth analysis, and to provide efficient and fast operation control means, so as to solve the following problems faced by the current dispatching and operation automation system of power system:

1) Under the background of the expanding scale of power grid, the data information explosion problem faced by dispatching operation is becoming more and more serious. With the acceleration of power system construction and the increasing scale of power grid, unprecedented challenges have been brought to the data monitoring and management of dispatchers and operators. On the one hand, the number of substations, transmission lines and other equipment is expanding, which makes the scale of data information expand; on the other hand, with the increasingly complex operation mode of AC-DC hybrid system in South Network, the difficulty of operation analysis such as transient stability and dynamic stability is increasing. These changes make the dispatch operators face increasingly serious data explosion problem.

2) Under the background of the continuous improvement of the level of power system dispatching automation system, information exchange between different automation systems is a problem. Under the background of expanding scale of power grid ,in order to meet the requirement of dispatching and operation command of power system under the background of expanding scale of power grid, a large number of technical research and tackling key problems have been carried out by major power grid companies in recent years, and many technical support systems have been developed and implemented, including EMS system, generation planning system, load forecasting system, safety checking system, etc. Most of the above-mentioned automation systems focus on one aspect of power system operation. The lack of effective integration and deep mining of a large amount of data information makes it difficult for dispatchers to obtain the overall picture of the power system through a platform, let alone to further analyze a large number of integrated and combed data in order to enhance their operational analysis and decision-making ability.

3) Under the background of the continuous improvement of the lean requirements of power system operation control, the practical needs of efficient decision-making analysis for dispatchers and operators are put forward. Lean operation of power system requires dispatchers to have stronger operation control ability to ensure safe, stable and economic operation of power system. However, facing huge data information and scattered and complicated decision-making system, dispatchers and operators are often at a loss.

III. Key Technologies for Operating Cockpit

A. Technical Architecture

As the front end of OS2 system, the operating cockpit is a platform connecting OS2 system and dispatching operators. Indicators are the core of operating cockpit in power system,data analysis and data mining of indexes are important technologies in operating cockpit. Combining the actual experience of OS2 system development and construction, the core technology system of the operating cockpit is shown in Figure 3, including data processing, data analysis and data display.
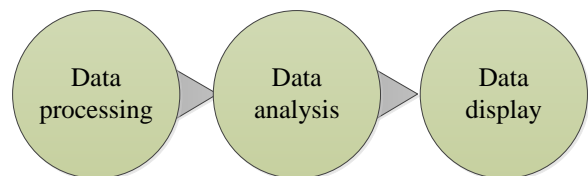
Figure 3. Indicator Processing Technology in Operating Cockpit

B. Data Processing Technology Based on Dimensionless Index

Data processing is the basic content of the technical framework for operating cockpit. Its purpose is to process the "raw" data with errors and inconsistent standards into "familiar" data through data screening, comparison and processing.

Power system has many data sources and complex structure, which can be divided into 500 kV, 220 kV and 110 kV in terms of voltage level, generation side, transmission side, substation side, distribution side and power side in terms of link, and primary and secondary information in terms of equipment classification. The huge and complex data system requires dispatchers to have a deep understanding of the meanings of different types of data, to be able to quickly identify their rationality, to determine whether they are within the qualified range, and to accurately grasp the distance from the qualified range boundary.

With the continuous expansion of power grid scale, the above requirements are becoming more and more difficult for power system dispatchers and operators. Therefore, it has become the key content of data processing to dimensionless

multi-source data with different dimensions by means of non-dimensional index.

The so-called dimensionless index refers to the normalization of the index values in different dimensions based on the analysis of the characteristics of the index values, so that they can be transformed into dimensionless data with the boundaries of 0-100, similar to the "test scores", so that dispatchers and operators do not need to study the connotation of the data, and can quickly judge the current operation of the power grid according to the dimensionless data.

According to the actual situation of power system operation, the operation indexes of power system can be divided into three categories: maximum, minimum and interval.

1）Maximum type

The so-called maximal index refers to that if the value of such index is greater than a given value, it will be qualified. The larger the value, the better the operation of the power grid will be. Then its dimensionless formula can be expressed as follows:

$$\bar{x} = \begin{cases} 0 & x < x_0 \\ 60 + 40 \times \dfrac{x - x_0}{MAX - x_0} & x \geq x_0 \end{cases} \tag{1}$$

As shown in Formula (1), x and $\bar{x}$ are the values of the index before and after dimensionless, x_0 is the given as the qualified threshold and MAX is the upper limit of the value of the given index. It can be seen that if the index value is lower than the threshold, the score will be 0, and if it is higher than the threshold, the score will be linearized and adjusted according to the value.

2）Minimum type

The so-called mini-index means that if the value of such index is less than a given value, it will be qualified. The smaller the value, the better the operation of the power grid will be. Then its dimensionless formula can be expressed as follows:

$$\bar{x} = \begin{cases} 0 & x > x_0 \\ 60 + 40 \times \dfrac{x_0 - x}{x_0 - MIN} & x \leq x_0 \end{cases} \tag{2}$$

As shown in Formula (2), x and $\bar{x}$ are the values of the index before and after dimensionless, x_0 is the given as the qualified threshold and MIN is the upper limit of the value of the given index. It can be seen that if the index value is higher than the threshold, the score will be 0, and if it is lower than the threshold, the score will be linearized and adjusted according to the value.

3）Interval type

The so-called interval-type index means that the value of this kind of index is qualified in a given interval, but it is not qualified beyond that interval. The better the interval is, the better the operation of the power grid is. Then its dimensionless formula can be expressed as follows:

$$\bar{x} = \begin{cases} 0 & x < x_1 \\ 100 - 40 \times \dfrac{x_0 - x}{x_0 - x_1} & x_1 \leq x < x_0 \\ 100 - 40 \times \dfrac{x - x_0}{x_2 - x_0} & x_0 \leq x \leq x_2 \\ 0 & x > x_2 \end{cases} \tag{3}$$

As shown in Formula (3), x and $\bar{x}$ are the values of the index before and after dimensionless, x_2 and x_1 are the upper and lower limits of the interval, x_0 is the optimal value point in the interval, satisfying $x_1 < x_0 < x_2$. It is stipulated that when the value of the index is x_0, the value after dimensionless is 100.

C. *Data Analysis Technology with Multidimensional Index System as the Core*

The essence of multi-dimensional index system lies in the process of integrating and clustering similar indexes from different dimensions according to their connotative characteristics and extension needs.

At present, the multi-dimensional index system of power system can be divided into the following dimensions:

1) Subject dimension is based on the goal of power grid operation, which divides power grid operation indicators into four categories: safety, economy, energy saving and environmental protection;

2) The spatial dimension divides the grid operation indexes into five categories: generation side, transmission side, substation side, distribution side and power side according to the position of the corresponding equipment in the production and operation of the power system;
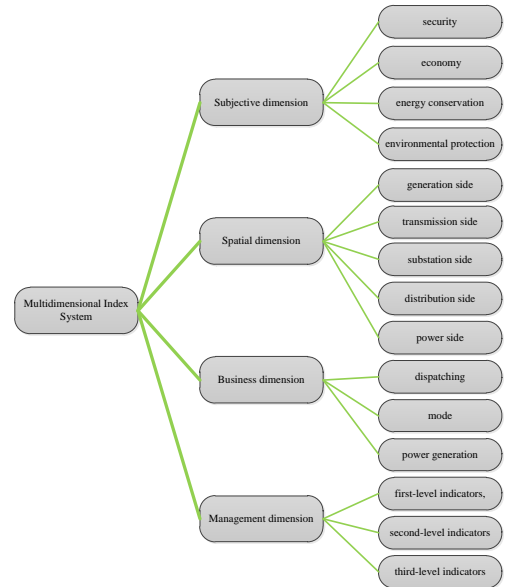

Figure 4. Multidimensional Index System

3)According to the division of business functions of different specialties in power system dispatching and operation, the operational indicators of power system are divided into dispatching, mode, power generation, water dispatching, automation, protection, communication and so on;

4) The management dimension divides the operation indicators into first-level indicators, second-level indicators, third-level indicators and basic indicators from the level of power system dispatching and operation management.

The above multi-dimensional index system can be represented as shown in Figure 4 below. It should be noted that the above dimensions are more integrated clustering of indicators based on business needs, which means that an indicator may be subordinate to multiple plates at the same time in a certain division angle, such as in the business dimension, coal consumption indicators are not only dispatching indicators, but also generation indicators.

D. Data Presentation Technology Based on Data Visualization

The so-called data visualization refers to in-depth analysis of individualized characteristics of index data in the process of data display, and selection of visualized display mode based on characteristics, so that dispatching operators can grasp the changing rule of data more accurately and efficiently.

At present, chart types commonly used in data visualization include graph, pie chart, waterfall chart, radar chart and histogram. The internal meaning of common data indicators in power systems and the applicable chart types are shown in Table 1.

TABLEI. ELECTRIC POWER SYSTEM INDICATORS AND CHART COMPARISON TABLES

Data index	Features	Applicable charts	Attention item
Voltage eligibility rate	Daily or monthly statistics, focusing on reflecting eligibility and demonstrating trends	graph	Need to increase eligibility line
Frequency eligibility rate	Daily or monthly statistics, focusing on reflecting eligibility and demonstrating trends	histogram	Because the pass rate of frequency is almost 100%, the bar chart is more suitable than the straight curve.
Accuracy of load forecasting	Daily or monthly statistics, focusing on reflecting eligibility and demonstrating trends	graph	Need to increase eligibility line
Installation structure	It is necessary to show the proportion or capacity of different types of power supply or different areas of power supply.	waterfall chart	Because pie charts are more suitable for displaying the proportion relationship, the total amount can not be displayed, and waterfall charts are more suitable.
composite index	Need to show coordination between different indices	radar chart	
Generation capacity	Demonstrate the changing trend and the proportion structure of different types of power supply/different groups of power supply	radar chart	
Power supply	Demonstrate the proportional structure of each region	pie chart	
Electricity sales	Demonstrate the proportional structure of each region	pie chart	
Regional power supply	The emphasis is on displaying the proportional relationship.	pie chart	

IV. SYSTEM DEVELOPMENT OF OPERATING COCKPIT

A. Three-tier Network-Provincial-Local Architecture

The cockpit of power system operation needs to cover three-level dispatching agencies of network, provinces and prefectures. According to the six unified construction principles, the whole OS2 system needs to be connected up and down, and its three-tier network-provincial-territorial architecture is shown in Figure 5. Under this architecture, the three-tier main stations of network, province and region are arranged separately, and the data connection is realized by vertical OSB. The operating cockpit is developed according to the hierarchy. The upper dispatching agency can access the lower dispatching agency cockpit online through the vertical OSB and the horizontal OSB.
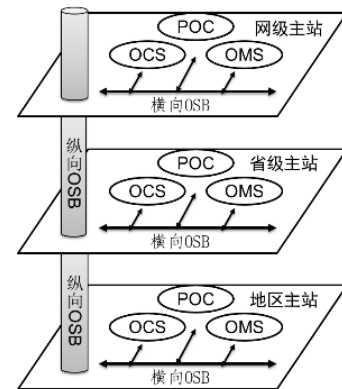
Figure5. Third-level structure diagram

B. System Function Configuration

According to the actual needs of power system dispatching and operation, the structure of the automatic system for operating cockpit is shown in Fig. 6. OS2 system data center is deployed in Secure II and Secure III. It integrates the data related to power system operation as data information support. The integrated data comes from EMS system, WAMS system, water dispatching system, credit guarantee system and so on. The main station of the operating cockpit is deployed in Secure II, and the required data are collected from the data center located in Secure II for statistical analysis and calculation. At the same time, through the forward isolation device between Secure II and Secure III, WEB version display server and mobile terminal display server for query only are set up in Secure III to meet the needs of mobile access analysis and computer office platform access analysis.
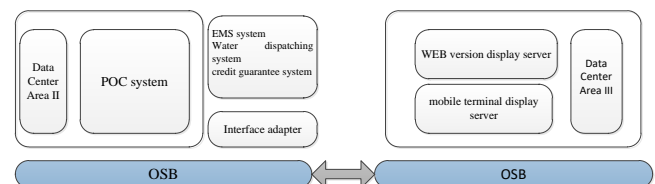
Figure 6. Figure of automation system architecture

C. system Development

On the system development, in order to meet the requirements of different operating systems, the cross-platform development technology of Java language is mainly used to meet the needs of different dispatching agencies at three levels of the South Network.

The display interface of the two-level Network-Provincial operation cockpit is shown in Figs. 7 and 8.

Figure 7. General adjustment operation cockpit

Figure 8. middle run cockpit

V Operation Benefits of the System

1) Improve the automation level of power grid dispatching and operation, and solve the problems of multi-source and inconsistent data. Operating cockpit, as an important part of OS2 system, plays a bridge role between dispatching operators and automation system. Especially, OS2 system realizes data center, solves a series of problems such as multi-source and non-uniformity of data, and effectively improves the automation level of power grid dispatching and operation.

2) Improve the analysis level of power grid dispatching operation command link, and provide a comprehensive and objective analysis basis for dispatchers. For a long time, dispatchers have limited a large amount of data and information in power grid automation system to safety type data, lacking of means for multi-dimensional benefit analysis of power grid operation economy, energy saving, environmental protection and so on. The multi-dimensional index system of operation cockpit objectively shows the change of multi-dimensional objectives of power grid, especially reveals the substitution and coordination characteristics of multi-dimensional objectives. As the decision-making basis of power grid dispatchers, comprehensive comparison of operation schemes plays a very important role in improving the comprehensive efficiency of power grid operation.

3) The level of power grid dispatching operation management has been improved, and the benchmarking management among three-level dispatching agencies has been realized with the index as the grasp. Vertical management of provincial and prefectural dispatching agencies and business benchmarking management between dispatching agencies at the same level are difficult

management problems faced by China Southern Power Grid and other power grid companies. The construction of operation cockpit solves the key problem of power grid operation index, divides the focus of attention of dispatching agencies at different levels from the management dimension, forms the first, second and third-level indicators, meets the management requirements of different dispatching agencies, and realizes the peer-to-peer benchmarking between dispatching agencies.

In summarizing the above successful experience, we must also see the existing problems in the process of popularization and application of the operating cockpit, including the low actual utilization rate of the cockpit at the county level, and the lack of cockpit function in some areas due to data problems. The fundamental reason is that before the construction of integration, the dispatching business processes and data base are quite different, which makes it difficult for top-level design to land in the process of integration. For this reason, China Southern Power Grid has adopted a variety of technical means to solve the above problems. On the one hand, business process standardization, relying on business guidance, combing and clarifying the key business and function division of different specialties in different regions in an all-round way to ensure the consistency of business processes; on the other hand, increasing investment in basic equipment, salvaging the technical basis, especially for the first county with great differences in technical strength. Level I power supply enterprises should increase their business investment and comprehensively improve data quality.

Conclusion

It has been nearly seven years since the concept of operating cockpit was put forward, and it has been popularized and applied in three-level dispatching agencies of Southern Power grid, provincial and local governments. This paper systematically introduces the function orientation, key supporting technology, system development framework and system application benefits of the operating cockpit, and summarizes the outstanding functions of the operating cockpit in three aspects: upgrading the level of the dispatching and operation automation system, enhancing the operational command ability of the dispatching and operation, and enhancing the operational management business of the dispatching and operation. It will have a very important practical significance for the technical promotion of operating cockpit.

References

[1] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989. J. Wang, H. Zhou, "Conceptual Design and the Future Development for Operation Smart System in China Southern Power Grid", *IEEE Transactions on Smart Grid*, vol. 4, pp. 1621-1629, 2013.

[2] J. F. Wang, X. C. Wu, H. F. Zhou, "Research and development of Operation Smart System in CSG," in *San Diego, CA. 2012 IEEE Power and Energy Society General Meeting.*, pp. 1-6.

[3] Sang-Seung Lee, Min-Uk Yang, Kern-Joong Kim, etc al. "Northeast Asia power system interconnection and smart grid operation strategies in South Korea," in *Vncouver, BC, Canada. 2013 IEEE Power & Energy Society General Meeting.*, pp. 1-5.

[4] ZHOU Zhemin, WU Wenchuan, DONG Xuzhu, et al. "A Preliminary Investigation on Smart Grid Operation Cockpit," in *Asia. 2012 IEEE Innovative Smart Grid Technologies Meeting.*, pp. 1-4.

[5] KANG Taifeng, GE Liang, TAN Zhihai, et al. "Investigation on KPI for self-healing smart distribution system operation cockpit,",in *China. 2013 IEEE International Conference of IEEE Region 10.*, pp. 1-3.

[6] LIU J, WEN B, LI P, et al. "Strategy-based Dispatching model and its Support System," in *Auckland, New Zealand. 2012 IEEE International Conference on Power System Technology (POWERCON).*, pp. 1-5.

[7] ZHENG Wenjue, YU Nanhua, CHEN Jiongcong, et al. " Implementation of IOC and its application in distribution system operation and maintenance management," in *Shanghai,China. 2012 China International Conference on Electricity Distribution.*, pp. 1-5.

[8] R. Wang, Y. Guo, F. Guan, J. Ma, S. Sun and Z. Yang, "A Construction Method of Power Grid Abnormal State KPI Based on AHP," *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*, Changsha, China, 2019, pp. 2498-2501.

[9] J. Zhang and J. Lan, "Research and application of large-scale renewable energy control & monitoring system key technologies," *8th Renewable Power Generation Conference (RPG 2019)*, Shanghai, China, 2019, pp. 1-7.

Study on Energy Harvesting of Open-close Current Transformer

Chi Tan
School of electrical engineering
Xi'an Jiaotong University
Xian, China
tanchi@stu.xjtu.edu.cn

Yanzhen Zhao*
School of electrical engineering
Xi'an Jiaotong University
Xian, China
zhaoyzh@mail.xjtu.edu.cn

Zijian Tang
School of electrical engineering
Xi'an Jiaotong University
Xian, China

Abstract—In order to solve the power supply problem of the online monitoring sensor of the high-voltage transmission line, this paper has conducted a study on the energy harvesting of an open-close current transformer (CT). According to the principle of CT induction, the relation between output power and resistor is derived. The silicon steel was selected as magnetic core material by calculating the effect of air-gap size on the initial permeability of different core materials. And a bleeder circuit is designed to solve the problem of CT saturation when the primary current is large. In order to verify the feasibility of the derivation, several simulation has been performed using Saber. Finally, the experiment device is set up and the results showed that the open-close CT could stably output power of 0.5 W when primary current is between 18.6 and 500 A, meeting the demand of sensor power supply.

Keywords—Open-close CT, magnetic core material, bleeder circuit, Saber simulation

I. Introduction

At present, many countries in the world are accelerating the construction of smart grid. Compared with traditional grid, smart grid has a large number of sensors. These sensors provide a wide range of real-time data, such as voltage, current, frequency, power quality, and temperature of different devices. And these data enable the staff to fully understand the operation status of the power grid, control the operation of power generation, transmission, distribution, etc, and comprehensively maintain the health and stability of the power grid [1,2].

The state grids have built a dense high-voltage transmission lines for power transmission in order to rationally allocate power resources. Generally, the high voltage transmission line has a wide distribution zone and remote position, which makes it difficult to obtain the conventional power supply for the monitoring equipment installed on the transmission line. Various methods of obtaining electrical energy appear for maintaining the normal operation of the sensor. The common energy harvesting methods of monitoring equipment for high-voltage transmission lines are solar energy harvesting [3], laser power supply [4], capacitance voltage divider power harvesting [5], mutual inductance power supply [6-9] etc. Solar energy harvesting cannot be used in a large area due to the uncertainly of daylight hours; The laser power supply has a short life span, high operation and maintenance costs; The capacitive voltage divider power harvesting cannot provide enough power for the load due to its limited energy. Compared with these energy harvesting methods, the energy harvesting CT of the mutual inductance power supply method has the advantages of simple structure, small

size, low cost, stability and reliability, and has been widely studied by many scholars in this field.

Compared with other energy harvesting CT, open-close CT is convenient to disassembly and installation on the high-voltage transmission line. Therefore, this paper studies the open-close CT. Actually, there are two problems for energy harvesting CT [6]: (1) CT devices immersed in saturation when the primary current is large, (2) CT devices cannot obtain enough energy to supply the load when the primary current is small. In this study, the output capacity of open-close CT at a small current on the primary side has been explored through the analysis on the best resistor and the relationship of air-gap size and magnetic core material. Then, a way of anti-saturation ability of open-close CT with large current on the primary side has been referred by connecting the bleeder circuit. Finally, the feasibility of the method is tested by simulation and experiment.

II. Principle and Design

A. The Principle of CT Power Supply

The conventional CT power supply model is shown in Fig. 1. i_1 is the current flowing through a high-voltage transmission line. R is load and i_2 represents the current flowing to the load.

When i_1 flows through transmission line, the magnetic core of CT generates alternating magnetic flux with the same frequency as the transmission line due to electromagnetic induction. And i_2 will be generated. A stable output source can be obtained by processing and converting the voltage of R.
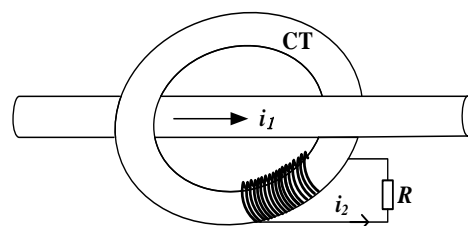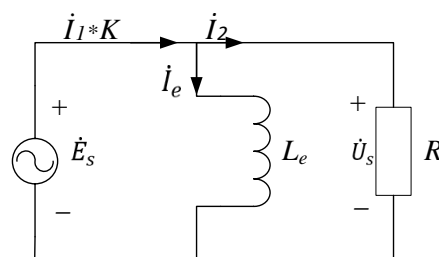

Fig. 1. CT power supply model

Fig. 2. The equivalent circuit of CT [10]

567

Ignoring the influence of leakage inductance, coil resistance and core loss on the circuit, the equivalent circuit of the ideal current transformer is shown in Fig. 2 [10]. Where $\dot{I}_1$ represents the phasor of primary current; K is the current turns ratio; $\dot{I}_e$ is the excitation current phasor; $\dot{I}_2$ represents the phasor of secondary current ; L_e is excitation inductance. The output voltage is

$$U_s = \left| K\dot{I}_1 \frac{j\omega L_e R}{j\omega L_e + R} \right| = K I_1 \frac{\omega L_e R}{\sqrt{R^2 + (\omega L_e)^2}} \quad (1)$$

The power of load is

$$P = \frac{U_s^2}{R} = (KI_1)^2 \frac{(\omega L_e)^2 R}{(\omega L_e)^2 + R^2} \quad (2)$$

Find the derivative of power

$$\frac{dP}{dR} = (KI_1\omega L_e)^2 \frac{(\omega L_e)^2 - R^2}{\left[(\omega L_e)^2 + R^2\right]^2} \quad (3)$$

The derivative is zero and *P* is maximized when

$$R = \omega L_e \quad (4)$$

Substituting (4) into (3), the maximum power delivered to R can be obtained as follows

$$P_{Max} = 0.5R(KI_1)^2 \quad (5)$$

The resistor was set as the best resistor at this point. That is, connecting the best resistor needs a small primary current when the same power is taken.

B. The Selection of Core Material

Compared with the material with low initial permeability, the material with high initial permeability have the advantages of large induction value with the same winding number and good low frequency characteristics. Its drawback is easy saturation under the large primary current. And table 1 shows the magnetic parameters of three common cores materials.

TABLE I.   THE MAGNETIC PARAMETERS OF THE THREE CORE MATERIALS

Magnetic parameter	Silicon stell	Permalloy	nanocrystalline
Saturation induction density (B_s, T)	2.1	0.7	1.2
Initial permeability(μ_i)	1500	100000	100000
Maximum permeability(μ_m)	40000	450000	400000
Relative permeability(μ_r)	7000~10000	20000~200000	1500~300000
lamination factor	0.95	0.9	0.7

it can be seen from the table 1 that μ_i of permalloy and nanocrystalline materials is very high, which can induce a relatively large output power when the primary current is small. The lamination factor and B_s of silicon steel sheet material is the largest, which can enhance the effective area of the core and the upper limit of large current on the

primary current respectively. The absolute smoothness of the cutting surface cannot be guaranteed because of the immature cutting process when the open-close CT core is cut, and the CT is cut against each other [7]. Therefore, an air-gap is inevitably introduced, which will lead to a decrease in the output power of the circuit at the same primary current and increase the lower limit of the primary current [11][12].

Suppose the air-gap size of open-close CT is δ. The relative permeability μ_{eq} of the core can be calculated by the following formula

$$\frac{l}{\mu_{eq}\mu_0} = \frac{l-\delta}{\mu_r\mu_0} + \frac{\delta}{\mu_0} \quad (6)$$

Where, *l* is the length of magnetic circuit; μ_0 is the Vacuum permeability. Simplify the equation

$$\mu_{eq} = \frac{1}{\frac{l-\delta}{\mu_r l} + \frac{\delta}{l}} \quad (7)$$

Take δ = 0.01 mm, *l* = 267 mm, μ_r = maximum as an example, and substitute into Equation(7). The μ_{eq} of the three magnetic materials is 7275.40, 23555.46 and 24517.98 respectively; When δ = 0.05mm, the equivalent initial permeability of the three magnetic materials is 3481.32, 5201.16 and 5246.63 respectively; When δ = 0.1mm, the equivalent initial permeability of the three magnetic materials is 2107.50, 2634.84, 2646.46. As can be seen from the above analysis μ_i of nanocrystalline and permalloy is no longer dominant when δ exceed 0.05mm. Considering that nanocrystalline and permalloy cores are difficult to be made into open structure due to the limitation of cutting process in practice. And the cutting process of silicon steel sheet core is more mature. The [13] proves that δ introduced by cutting silicon steel is about 0.02 mm, and the δ introduced by cutting nanocrystalline is at least 0.05 mm. Therefore, this paper selects CT made of silicon steel for research.

The magnetic resistance (R_m) of open-close CT is

$$R_m = \frac{l-\delta}{S\mu_0\mu_r} + \frac{\delta}{S\mu_0} \quad (8)$$

Where, S is the cross-sectional area of the CT. The relationship between R_m and magnetic resistance is

$$L_e = \frac{N_2^2}{R_m} \quad (9)$$

The best resistor value can be roughly estimated by calculating (4), (8) and (9).

C. Bleeder Circuit

The bleeder circuit can not only prevent CT saturation, but also stabilize the output voltage within a certain amplitude, which is convenient for the back circuit processing. At the same time, the adaptive feature of the bleeder circuit can reduce the ripple of filtered circuit. The bleeder circuit show in Fig. 3. Where, D_1 is diode to prevent current backflow; D_2 is the regulator diode; D_3 is thyristor; C is filter capacitor. R_1 is a resistor whose voltage can turn on thyristor.
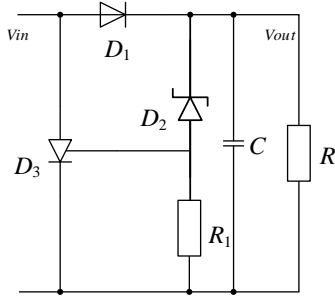
Fig. 3. The bleeder circuit

Its working principle is as follows: When the induced electromotive force on the secondary side is rectified by the bridge rectifier circuit. C is charged, and the voltage across C rises; when the voltage of C rises to a certain value, D_2 is broken down. The voltage of R_1 increase until the gate trigger voltage of D_3 is reached. Then, D_3 conducts, and the open-close CT is shorted. C discharge supplies R, which lead to decrease of output voltage. When the instantaneous value of output voltage of bridge rectifier circuit reaches 0, D_3 does not conduct since the anode current is less than the minimum value to maintain conduction. And the whole process repeats itself.
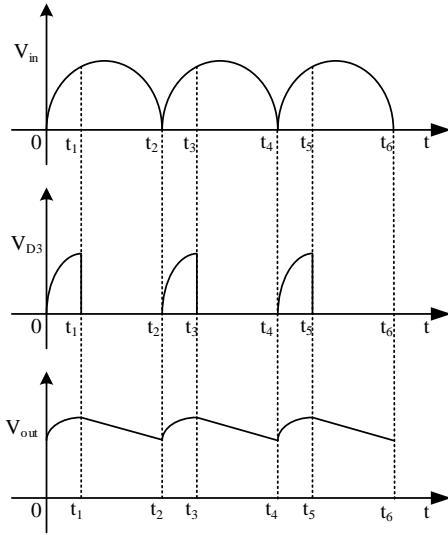
Fig. 4. The voltage relationship

the voltage relationships between V_{in}, V_{D3} and V_{out} are shown in Fig. 4. $0 \sim t_1$ is the charge and $t_{1\sim} t_2$ is the discharge within a period. The effective value of the final output voltage is

$$V_{out} = \frac{1}{\pi} \int_0^{t_1} V_{in} \sin(\omega t) d(\omega t)$$
$$0 < t_1 \leq \frac{\pi}{\omega} \tag{10}$$

III. SIMULATION AND ANALYSIS

In this paper, the parameters of the open-close CT are shown in Table 2. The magnetic core is made of silicon steel sheet. Considering that the general sensor power is about 0.5W, there has loss in the DC-DC circuit. So, 0.6W is taken as the lower energy limit.

TABLE II. OPEN-CLOSE CT PARAMETER

Core parameter	value
inner diameter （mm）	57
external diameter （mm）	104
turns ratio	1:200
height （mm）	32
air-gap size （mm）	0.01

A. Simulation Model

Saber simulation software is an EDA software from Synopsys in the United States, which is known as the world's most advanced system simulation software. And the magnet component tool (MCT) of the Saber software can establish accurately CT model that can precisely simulate the actual situation of CT and solve a series of problem including the design of back circuit and the choice of components.

The core parameters in Table 1 and Table 2 are entered to build an open-close CT simulation model by using the MCT of the Saber software. The simulation circuit diagram is shown in Fig. 5:
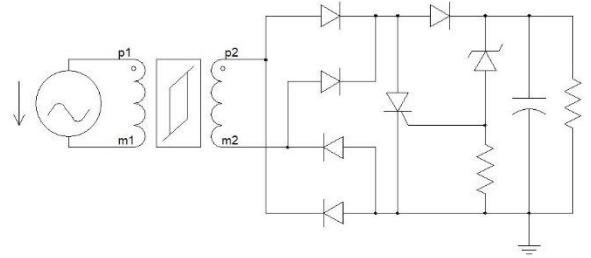
Fig. 5. Simulation model

The Fig. 5 hasn't DC-DC circuit because the LM2576 doesn't exist on the library of Saber and there hasn't the Saber model of LM2576.
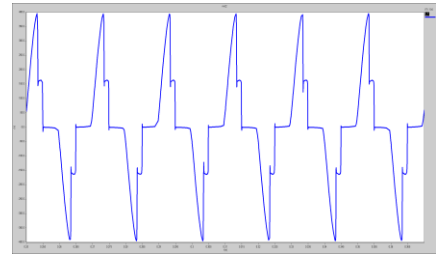
Fig. 6. the load output waveform when the primary current is 200 A

Only add a load on the secondary side, set the load value to 1000 Ω, the size of primary current frequency is 50 Hz and the value is 200 A. The secondary side output voltage waveform diagram is displayed on Fig. 6. As can be seen from the figure, the output voltage waveform of the circuit is seriously deformed at this time, and The CT has been deeply saturated.

B. Bleeder Circuit Test

According to (4), (8) and (9), the best resistor is calculated to be in the range of 518~675 Ω. The resistor should be reduced by 0.81 times since the resistor plus the rectifier filter circuit is equivalent to a load [14].So, the best resistor should be in the range of 632~833 Ω after adding the rectifier filter circuit.

The stabilizing voltage value of the regulator tube was set to 36 V and the primary current input is 17.5 A, 50 Hz. And the simulation model of Fig. 5 is simulated. The relationship between resistor and output power by changing the value of resistance is shown in Fig. 7. It can be seen from the figure that the circuit can output the maximum power when the resistor is about 700 Ω, which can reduces the lower limit of the primary current. The result is consistent with the previous analysis.
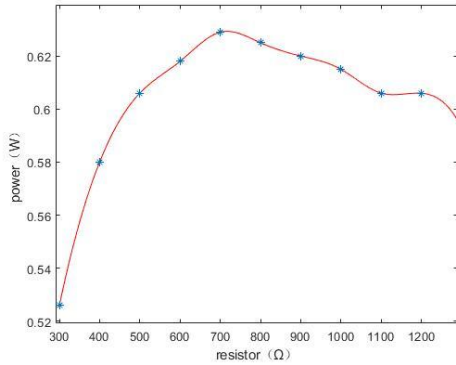
Fig. 7.   Output at different loads when primary current is 17.5 A

The resistance is set as 1000Ω for the convenience of field experiment and the different primary value are inputted to get the output waveform of Fig. 8 and Fig. 9.
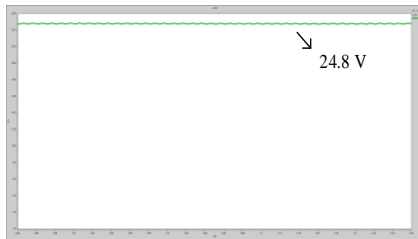
Fig. 8.   Output when primary current is 17.5 A

As shown in Fig. 8, when the primary current is 17.5 A, the output power is 0.615 W, which meets the requirements of the circuit. According to the output waveform in Fig. 9, the bleeder circuit works normally when the primary current is high. The output of the voltage is stable at 36.9 V. And the output voltage's ripple of Fig. 8 and 9 isn't exceeding 0.5%, which proves that the bleeder circuit can effectively decrease the ripple of output load.
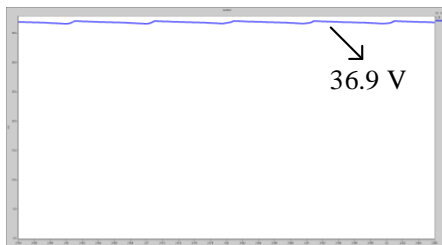
Fig. 9.   Output when primary current is 500 A

IV.   THE RESULT OF EXPERIMENT

In this experiment, the power supply generated by the strong current generator simulates the current of the high voltage transmission line. The sensors is replaced by resistor, and the whole experimental setup is shown in Fig. 10. First of all, DC-DC circuit is instead of resistor in order to verify the accuracy of the simulation experiment and the value of resistor is 985 Ω.
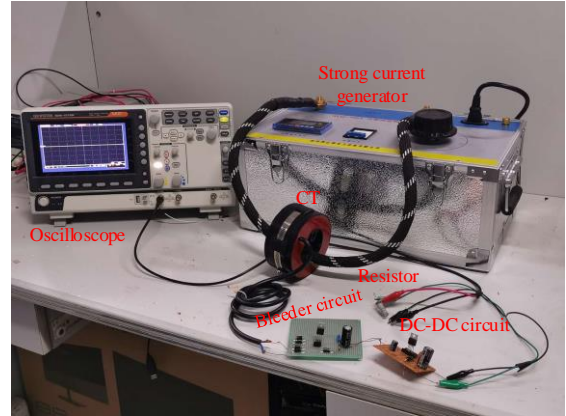
Fig. 10. Experimental facility
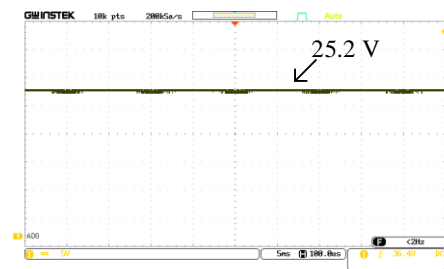
Fig. 11. The output voltage when the primary current is 17 A

Fig. 12. The output voltage when the primary current is 200 A
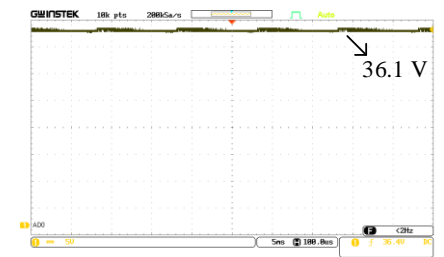
Fig. 13. The output voltage when the primary current is 500 A

Connect the circuit and adjust the primary current value to 17A, 200A and 500A to obtain the corresponding output voltage image, as shown in Fig. 11, 12, and 13. The power of 17A at the primary side was calculated, and the power value was 0.645W, which was almost equaled with the simulation experiment. From the output images of primary current value 200A and 500A, the Rms of output voltage is almost the same.

The DC-DC circuit built by LM2576 is selected and connected to the circuit. Table 3 shows the different load corresponding to the power. And Fig. 14 is output voltage

when R=49.9 Ω. Neglecting the error, it can be concluded that the output power of circuit is 0.5 W when the primary current is18.6 A.

TABLE III.   OUTPUT VOLTAGE AND POWER AT DIFFERENT LOADS

Resistor (Ω)	Primary current （A）	Output voltage (V)	Power (W)
48.9	18.6	4.88±0.08	0.487
49.5	18.6	4.88±0.08	0.481
49.9	18.6	4.88±0.08	0.477
50.7	18.6	5.00±0.04	0.493
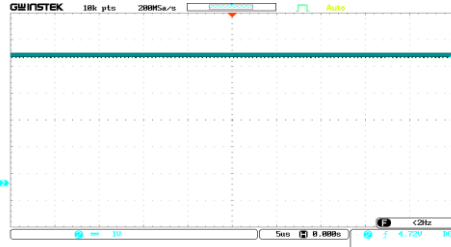
Fig. 14. The output voltage when R = 49.9 Ω

There has difference between 18.6 A and 17 A because of the fact that the back circuit is equivalent to a resistor when the DC-DC circuit isn't added. Actually, the DC-DC circuit and output load equivalent to front circuit may be a nonlinear resistor, not a pure resistor. And this resistor and output load aren't equivalent. Therefore, it's normal to have some errors.

TABLE IV.   OUTPUT VOLTAGE FOR DIFFERENT PRIMARY CURRENTS WHEN R=49.9Ω.

Primary current （A）	Output voltage(V)
18.6	4.88±0.08
100	4.88±0.08
200	4.88±0.08
500	4.88±0.08

From the Table 4, the output voltage is changeless when primary current increase. That is to say, the output voltage of open-close CT is still steady with the increase of primary current.

V. CONCLUSION

In this paper, the energy harvesting of open-close CT is studied. Two problems for energy harvesting CT is solved through the rational collocation of external circuit, and the stable power output is achieved.

1) The best resistor is found through studying the equivalent circuit of CT. It can maximize the output power under the same conditions, which can decrease the lower limit of primary current.

2) The influence of air-gap size on the relative permeability of magnetic core material is discussed. Combined with the difficulty and accuracy of cutting process and the influence of core parameters on energy harvesting, the silicon steel material is finally chosen as the magnetic core material for open-close CT.

3) The original CT was heavily saturated at a primary current of 200 A, but it is not saturated at a primary current of 500 A by connecting the bleeder circuit. And The upper limit of primary current input is raised.

4) Simulation and experiment verify the correctness of the above analysis. The last experiment proves that the open-close CT can output stable power 0.5W at primary current 18.6-500A, which is enough to sensor.

REFERENCES

[1] R. Morello, S. C. Mukhopadhyay, Z. Liu, D. Slomovitz and S. R. Samantaray, "Advances on Sensing Technologies for Smart Cities and Power Grids: A Review," in IEEE Sensors Journal, vol. 17, no. 23, pp. 7596-7610, 1 Dec.1, 2017, doi: 10.1109/JSEN.2017.2735539.

[2] Karimulla S , Ravi K . A Review on Importance of Smart Grid in Electrical Power System[C]// 2019 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC). 2019.

[3] H. Sharma, A. Haque and Z. A. Jaffery, "An Efficient Solar Energy Harvesting System for Wireless Sensor Nodes," 2018 2nd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 2018, pp. 461-464, doi: 10.1109/ICPEICES.2018.8897434.

[4] N. Wang et al., "One-to-Multipoint Laser Remote Power Supply System for Wireless Sensor Networks," in IEEE Sensors Journal, vol. 12, no. 2, pp. 389-396, Feb. 2012, doi: 10.1109/JSEN.2011.2161465.

[5] J. Zhang, P. Li, Y. Wen, F. Zhang and C. Yang, "A Management Circuit with Upconversion Oscillation Technology for Electric-Field Energy Harvesting," in IEEE Transactions on Power Electronics, vol. 31, no. 8, pp. 5515-5523, Aug. 2016, doi: 10.1109/TPEL.2015.2491960.

[6] Xiong Lan, He Youzhong, Song Daojun, et a1. Design on power supply for the transmission line on line monitoring equipment [J]. High Vollage Engineering, 2010. 36(9)：2252—2256.

[7] W. Zhou, Z. Liu, Q. Huang, Y. Jiang and Z. Cong, "Design of Magnetic Cores for Current Transformer energy harvesting devices," 2019 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Macao, Macao, 2019, pp. 1-5, doi: 10.1109/APPEEC45492.2019.8994809.

[8] D. Yuan, C. Kai, L. Jie-dong, C. Yang-chun, N. Ming and W. Wei, "Design of induced power apparatus for monitoring system of dynamic increasing capacity of high-voltage transmission lines," 2012 IEEE International Conference on Condition Monitoring and Diagnosis, Bali, 2012, pp. 774-777, doi: 10.1109/CMD.2012.6416261.

[9] Z. He, S. Nie, G. Qu, Y. Liu, G. Sheng and X. Jiang, "The Design of CT Energy Harvesting Power Supply Based on Phase-Controlled Method," 2012 Asia-Pacific Power and Energy Engineering Conference, Shanghai, 2012, pp. 1-5, doi: 10.1109/APPEEC.2012.6307318.

[10] J. Ahola et al., "Design considerations for current transformer based energy harvesting for electronics attached to electric motor," 2008 International Symposium on Power Electronics, Electrical Drives, Automation and Motion, Ischia, 2008, pp. 901-905, doi: 10.1109/SPEEDHAM.2008.4581191.

[11] D. Muthumuni, P. G. McLaren, W. Chandrasena and A. Parker, "Simulation model of an air gapped current transformer," 2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.01CH37194), Columbus, OH, USA, 2001, pp. 705-709 vol.2, doi: 10.1109/PESW.2001.916940.

[12] Ferreira R S D A , Araujo J F D , Andrade F L M , et al. Influence of electromagnetic forces in the gaps of a protective CT[J]. Science, Measurement & Technology, IET, 2018, 12(7):872-877.

[13] Yue Tianchen, Liu Yadong, He Zhimin, et al. Power output characteristics of magnetic core in CT energy harvesting devices [J]. High Voltage Apparatus, 2015, 51(1): 18-23

[14] R. Wang, Z. Wang, J. Du, J. Wu and X. He, "Design considerations of maximum energy harvesting and voltage control from high voltage power cables," 2014 International Power Electronics and Application Conference and Exposition, Shanghai, 2014, pp. 920-923, doi: 10.1109/PEAC.2014.7037982.

Mining Association Rules of Distribution Network Equipment Based on Genetic Algorithm

Wenxiong Mo
Guangdong Power Grid Co., Ltd
Guangzhou power supply
Guangzhou, China
gzmwx@139.com

Zhong Xu
Guangdong Power Grid Co., Ltd
Guangzhou power supply
Guangzhou, China
348867958@qq.com

Simin Luo*
Guangdong Power Grid Co., Ltd
Guangzhou power supply
Guangzhou, China
81157183@qq.com

Chao Chen
Shaanxi Key Laboratory of Smart Grid
& the State Key Laboratory of
Electrical Insulation and Power
Equipment
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
610825983@qq.com

Yiming Kong
Shaanxi Key Laboratory of Smart Grid
& the State Key Laboratory of
Electrical Insulation and Power
Equipment
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
1063560817@qq.com

Xuanda Lai
Shaanxi Key Laboratory of Smart Grid
& the State Key Laboratory of
Electrical Insulation and Power
Equipment
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
lxd593@stu.xjtu.edu.cn

Abstract—**Due to the wide variety of 10kV distribution network equipment and the complex operating environment, various defects have emerged frequently. Because the distribution network is directly connected to users, the distribution network equipment defects will lead to the occurrence of distribution network failures if not processed in time, affecting the safe and stable operation of the power grid. In this paper, a genetic algorithm(GA)-based method is used to find useful association rules from the defect data set of distribution network equipment. In this method, genetic algorithm is used to avoid scanning transaction database multiple times and improve the efficiency of mining association rules. Moreover, the confidence factor is used to objectively measure the interest of the association rules without providing a minimum support threshold. In this paper, the defect association analysis of the main distribution equipment in a region is carried out, and the association results are obtained, which can help the maintenance personnel to find the family defects of the equipment and speculate the type of defects, providing support for the maintenance and repair of equipment.**

Keywords—distribution equipment, defect association rules, genetic algorithm

I. INTRODUCTION

The main equipment of the distribution network, including distribution transformers, switchgear, overhead lines, cables, etc., is the final link to supply power to customers, and plays an important role in the power network. Due to its wide variety and complex operating environment, the frequent occurrence of defects in power distribution equipment makes it a weak part of the power system, so it is very important to analyze the defects of distribution equipment. During the long-term operation of the distribution network equipment, a large amount of defect data is recorded and accumulated through inspection and testing. On the one hand, these data are only used to guide the defect elimination of equipment and perform simple statistical analysis and chart display, without further mining the useful information hidden in the data. On the other hand, the defect statistics work needs to be done manually, which is inefficient, and because of the huge amount of data, it is difficult for analysts to find the association rules between defect data in the form of charts. With the continuous improvement of power grid defect data management and the

rapid development of data mining technology, using association analysis in data mining to conduct association rule mining on a large number of defect data of distribution equipment can not only improve the utilization rate and analysis efficiency of defect data, but also discover the family defect of equipment through association results, speculate the defect type, and provide support for the maintenance personnel to identify the weak links of the distribution equipment and specify the inspection plan.

At present, data mining technology has been widely used in the power. The research on the defects of distribution equipment mainly focuses on the text classification of power distribution equipment defects [1] and the image recognition of power equipment defects [2]. However, most of the research on defect analysis of distribution network equipment remains at the level of experimental analysis of specific equipment. In practical applications, defect data are usually only used for simple statistical analysis, there are few studies on defect association analysis of the distribution network equipment and the analysis methods mainly include Apriori, FP-growth and their improvement algorithm [3-5], with low mining efficiency. Reference [6] used Apriori algorithm to analyze transformer defects, reference [7] based on improved Apriori algorithm to monitor the family defects of transformers, and an improved FP-growth algorithm was presented to analyze transformer defects [8]. As a widely used optimization algorithm, genetic algorithm is mainly used in the optimization of economic load dispatch [9], distribution network reconfiguration [10], power flow optimization [11] and other aspects of the power system, but it is not effectively applied in the defect data mining of distribution network equipment.

In this paper, the defect level, defect condition, defect type, equipment manufacturer, and equipment operating life in the defect record data are selected to constitute the defect data set for the association analysis of distribution equipment. After encoding the defect data set with real numbers, the selection, crossover, and mutation in the genetic algorithm are used to optimize the current rules to generate the next generation of rules. After multiple iterations, a set of more ideal rules will be obtained. This method obtains the association rules by specifying the minimum confidence and the lift. Compared with the traditional Apriori method, this method fully utilizes

the genetic algorithm to search for the global optimal solution, improves the mining efficiency of association rules, and effectively finds the valuable rules hidden in the defect record data.

This paper is organized as follows. In section 2, the defect level, defect condition, defect type, equipment manufacturer, and equipment operating life in the defect record data are selected to constitute the defect data set for the association analysis of distribution equipment. The basic concepts of association analysis and the genetic algorithm are briefly introduced in section 3. In section 4, a genetic algorithm-based method using credibility factor as fitness function is proposed to find useful association rules from the defect data set of distribution network equipment. In section 5, the defect association rules of the main distribution equipment in a region is obtained and discussed. Meanwhile, the operating efficiency of the method is also evaluated. Finally, some conclusions are shown in section 6.

II. DEFECT RECORD DATA SET

Defect management is one of the most important management of electric enterprise production technology [12]. In order to ensure the health level of the equipment in the power system, maintenance personnel need to record and report equipment defects during the daily inspection process. At present, a large amount of historical defect data of distribution equipment is stored in the defect management system, which provides a data basis for the association analysis of related equipment.

Defect records generally include four categories of basic equipment information, substation information, basic defect information and defect handling information, each of which contains multiple attributes. In this paper, the defect level, defect condition, defect type, equipment manufacturer, and equipment operating life in the defect record data are selected to constitute the defect data set for the association analysis of distribution equipment according to the data quality and actual application. Taking the defect record data of the distribution transformer as an example, Table 1 shows the details of each attribute record in the defect data set, in which the defects and equipment manufacturers are only partially displayed due to the large number.

TABLE I. DEFECT RECORD DATA SET

Defect Level	Defect Condition	Defect Type	Manufacturer	Operating Life (years)
General	Oil leakage	Bad condition	manufacturer 1	[0,10]
Significant	Porcelain damage	Insulation problem	manufacturer 2	[10,20]
Emergency	Oil discoloration	Equipment damaged	manufacturer 3	[20,30]
/	Silicone discoloration	Abnormal protection device or meter	manufacturer 4	More than 30
/	…	External force failure	…	/

The defect level can be divided into three categories of "general", "significant" and "emergency" according to the severity of the defect. The defect condition is a simplified defect description of the equipment based on the defect description filled in by the maintenance personnel. The defect

type is the summary of the equipment defect. The equipment manufacturers are obtained from the equipment factory information, due to data confidentiality, the equipment manufacturers are classified as manufacturer 1, manufacturer 2 and so on in this paper. The equipment operation life is obtained by subtracting the defect occurrence time and the equipment operation time, divided by 10 years as the interval, and the equipment operating for more than 30 years is determined as the old equipment according to the regulations.

III. BASIC CONCEPTS

A. Association Analysis

The problem of association rule mining was proposed by Agrawal and others in 1993[13]. Its initial research goal was to discover knowledge about customer purchase behavior from the transaction database and express the discovered knowledge in the form of association rules. At present, association rules are widely used in the field of data mining.

$I = \{i_1, i_2, \cdots, i_n\}$ represents the set of all items in the data set, each i in the set is called an item, and the set containing multiple items in I is called the itemset. $T = \{t_1, t_2, \cdots, t_d\}$ represents a collection of all transactions. In this paper, T consists of the defect record data set, where each transaction record is a subset of a data item set I. The support of itemset A on T is the proportion of transactions containing A in T, which is expressed by (1).

$$sup(A) = P(A) = \left|\{t_i \mid A \subseteq t_i, t_i \subseteq T\}\right| / |T| \qquad (1)$$

Where $\left|\{t_i \mid A \subseteq t_i, t_i \subseteq T\}\right|$ represents the number of transactions containing A, also known as the support count of itemset A. $|T|$ represents the number of transactions contained in T.

The association rules are represented by the form $A \rightarrow B$, where $A, B \subseteq I$ and $A \cap B = \varnothing$. The support of the association rules is given by (2).

$$sup(A \rightarrow B) = sup(A \cup B) = P(A \cup B) \qquad (2)$$

The confidence of the association rule is defined as the ratio of the number of transactions containing itemsets A and B to the number of transactions containing A, given by (3).

$$conf(A \rightarrow B) = sup(A \rightarrow B) / sup(A) = P(B \mid A) \quad (3)$$

The support and confidence of the association rules respectively reflect the usefulness and certainty of the discovered rules. When mining association rules, the minimum support threshold and the minimum confidence threshold usually need to be given in advance, expressed as minsup and minconf respectively. The itemsets meeting the minimum support threshold are called frequent itemsets, and the rules that meet the minimum confidence threshold extracting from all frequent itemsets are called strong association rules.

B. Genetic Algorithm

Genetic algorithm were introduced by John Holland in the early 1970s [14]. Genetic algorithm is a random search algorithm that draws on the natural selection and genetic

mechanism. By simulating the biological evolution process, the problem to be solved approaches the optimal solution step by step from the random initial solution.

The genetic algorithm takes all individuals in a population as objects and uses randomization techniques to guide the efficient search of a coded parameter space. Among them, selection, crossover and mutation constitute the genetic operation of the genetic algorithm. The five elements of parameter coding, initial population setting, fitness function design, genetic operation design, and control parameter setting constitute the core content of the genetic algorithm.

IV. ASSOCIATION ANALYSIS BASED ON GENETIC ALGORITHM

According to the brief introduction of association rules in section 3, the goal of association rule mining is to obtain strong association rules. Traditional association rule analysis methods such as Aprior need to first obtain all frequent itemsets according to a given transaction database, then mine association rules and delete rules that do not meet the minimum confidence threshold, the rest are the strong association rules required by the user. The process of finding frequent itemsets can be regarded as a global search problem. The combination of genetic algorithm and association rule mining can effectively improve the efficiency of mining association rules for distribution equipment defects.

First encode the defect data set with real numbers. Next, define the fitness function of the genetic algorithm and select individuals with high fitness based on selection probability. Then, according to the given crossover and mutation probability, the current rules is genetically optimized to generate the next generation of rules. When the number of iterations reaches the maximum number given by the user, the algorithm stops and a set of ideal association rules is obtained. The following describes the association analysis steps in detail.
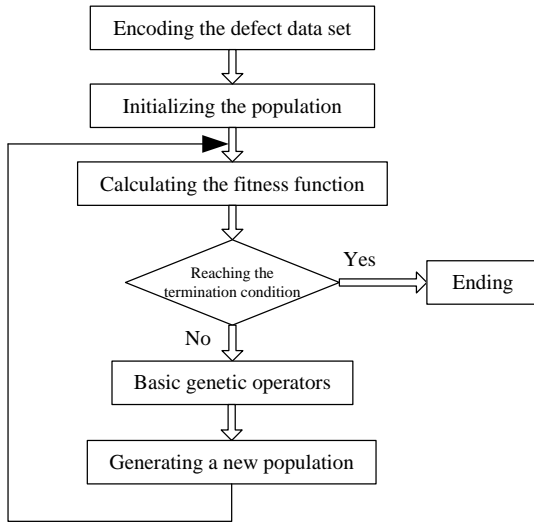

Fig 1. Association analysis based on genetic algorithm.

A. Encoding

Due to the large amount of defect data in the distribution network equipment and the attribute is classified attribute, the real number encoding is used in this paper, which is convenient for large space search. According to the arrangement of the attributes of the defect record data set in Table 1, all items in the data set are represented by real

numbers from 1. If the defect data set has j items in total, the set of all items can be expressed as $I = 1, 2, \dots, j$.

B. Fitness Function

As the link between genetic algorithm and association rule mining, fitness function is used to evaluate the quality of the generated rules and is related to the selection operation. In order to generate interesting rules required by users, a credibility factor is used to objectively measure the interest of the association rules and serve as the fitness function of the genetic algorithm. For association rules of the form $A \rightarrow B$, the fitness function $F(A \rightarrow B)$ is defined as follows:

$$F(A \rightarrow B) = \frac{sup(A \cup B) - sup(A)sup(B)}{sup(A)(1 - sup(B))} \qquad (4)$$

C. Selection

The selection operation is determined jointly by the fitness function value F and the selection probability sp, the selection probability is given by the user. The selection operator s is given by:

$$s(A \rightarrow B) = F(A \rightarrow B) \times sp \qquad (5)$$

All the generated association rules are selected with a probability of s, that is, the association rule with the larger fitness value is more likel y to be selected. All the selected rules are inherited to the next generation through crossover and mutation operations, which ensures the superiority of the rules of future generations.

D. Crossover

Crossover is the main method used to generate new rules to prevent the algorithm from falling into a local optimal solution. In this paper, two-point crossover is adopted. Under the premise of given crossover probability cp, two crossover points are randomly set in the encoding string, and then the genes in the two points are exchanged, thereby generating two new individuals.

E. Mutation

Mutation can not only improve the algorithm's ability to find the optimal solution, but also maintain population diversity and prevent premature convergence. In this paper, under the premise of the mutation probability mp given by the user, each attribute value of each rule is randomly mutated, and the mutated attribute value is different from the other attribute values of the rule.

V. RESULTS AND DISCUSSION

In this paper, the defect data of the key equipment of the distribution network from 2016 to 2019 in a regional power supply company are selected for defect association analysis. The key equipment of the distribution network includes four types of distribution transformers, switchgear, cable lines and overhead lines. First, perform data preprocessing on the original defect record data, delete the defect records with obvious errors or missing data, and select the defect level, defect condition, defect type, equipment manufacturer, and equipment operating life to constitute the defect data set for the associated analysis of distribution equipment , Which is

encoded in real numbers and used as an input data set for association rule mining.

A. Association Analysis Results

The association analysis based on genetic algorithm described in section 4 is used to mine association rules for the defect data set. Since five types of defect record attributes are selected for association analysis, the maximum length of association rules is five, and the specific parameter settings are shown in Table 2.

The association analysis of the defect data set is carried out according to different association rule lengths. Some association rules of the four key distribution network devices are shown in Table 3. Two association rules are taken as an example for each kind of equipment.

TABLE II.   PARAMETERS FOR GENETIC ALGORITHM

Parameter	Value	Description
minconf	0.6	the minimum confidence
lift	1	the minimum lift
popsize	1000	the size of population
cp	0.8	the crossover probability
mp	0.2	the mutation probability
sp	0.9	the selection probability
length	2, 3, 4, 5	the length of association rules

TABLE III.   SOME ASSOCIATION RULES

Equipment type	No.	Association Rule	Support Count	Confidence	Lift
Distribution transformer	1	{20,104} →{34}	293	1.0	1.5
Distribution transformer	2	{3,20,68} →{34}	21	1.0	1.5
switchgear	3	{2,151} →{36}	72	1.0	2.3
switchgear	4	{2,39,48} →{19}	29	0.71	2.5
cable lines	5	{3,18} →{20}	20	1.0	3.8
cable lines	6	{3,21,66} →{16}	20	1.0	3.9
overhead lines	7	{4,143} →{3}	220	0.75	1.5
overhead lines	8	{3,56,143} →{43}	32	0.78	1.2
…	…	…	…	…	…

By analyzing the defect association rules of the distribution equipment in Table 3, the following analysis results can be obtained: The association results can speculate the defect type and defect conditions, such as for defect rule 1, when the distribution transformer has been in operation for 10 to 20 years and the defect condition is oil leakage, the conclusion is the insulation problem, its confidence is 100% and the lift is 150%. When the antecedent is the switchgear manufacturer 3, the defect level is serious, and the defect type is insulation problem the confidence of the consequent of the rule is partial discharge exceeding the standard is 71%. In the process of equipment defect maintenance, the relevant staff can determine the possible defect types or defect conditions of the defective equipment according to the known conditions such as equipment defect conditions and equipment manufacturers, so as to improve the efficiency and accuracy of operation and maintenance.

At the same time, the association rules can find common defects in distribution equipment manufacturers, and provide support for the discovery of family defects in equipment. For example, for defect rule 6, cable manufacturer 5 has a strong correlation with the insulation damage of the intermediate joint, and the cable intermediate joint of this manufacturer is in the family defect list of the power supply bureau, which proves that the association rules are effective for family defect mining.

B. Time Complexity Comparison

In this paper, 1000, 2000, 5000, and 10000 transaction data in kosarak data sets were randomly selected as the comparison samples of algorithm time complexity. Traditional Apriori algorithm and the genetic algorithm used in this paper are used to mine association rules in different length data sets. The size of population in the genetic algorithm is set to 100 and 1000 respectively for operation. The computer configuration is Intel Core i5-6300HQ CPU @2.30GHz, and the memory is 12GB. The running time of different algorithms is shown in Fig. 2.
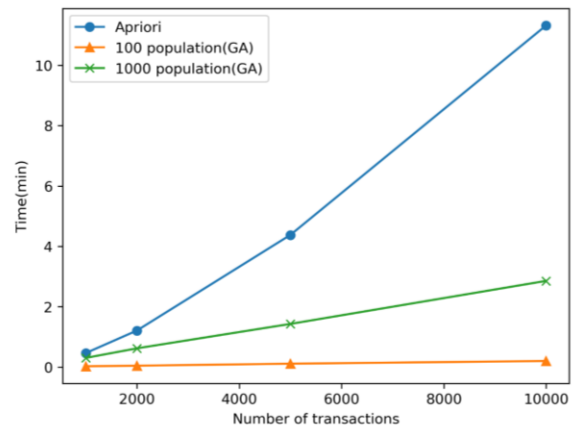
Fig. 2.   Time complexity comparison.

According to Fig. 2, when the length of the data set is the same, the running time of the genetic algorithm used in this paper is much shorter than the Apriori algorithm, and the size of population affects the efficiency of the genetic algorithm. As the length of the data set increases, the running time of the Apriori algorithm increases significantly, and the advantage of high efficiency of the genetic algorithm is obvious.

VI. CONCLUSION

In this paper, a GA-based method is used to find useful association rules from the defect data set of distribution network equipment. The experimental results show that the association rule mining efficiency of this algorithm is higher than that of traditional apriori, and there is no need for users to provide a minimum support threshold. The association results can help the maintenance personnel to speculate the possible defect types or defect conditions, find the family defects of the equipment, and improve the efficiency and accuracy of operation and maintenance.

REFERENCES

[1] J. Zhou, G. Luo, C. Hu and Y. Chen, "A classification model of power equipment defect texts based on convolutional neural network," International Conference on Artificial Intelligence and Security. Springer, Cham, 2019, pp. 475-487.

[2] M. Lan, Y. Zhang, L. Zhang and B. Du, "Defect Detection from UAV Images Based on Region-Based CNNs," 2018 IEEE International Conference on Data Mining Workshops (ICDMW). Singapore, 2018, pp. 385-390.

[3] Ansari, E. , Sadreddini, M. H. , Mirsadeghi, S. M. H. , Keshtkaran, M. , & Wallace, R. . "Tfi-apriori: using new encoding to optimize the apriori algorithm." Intelligent data analysis, 2018, 22.4, pp. 807-827.

[4] Sun, Li Na . "An improved apriori algorithm based on support weight matrix for data mining in transaction database." Journal of Ambient Intelligence and Humanized Computing, 2019.

[5] Li, M. , D. Ding , and Y. Yi . "Data Analysis of Tyre Quality Based on Improved FP-Growth Algorithm." Zhongguo Jixie Gongcheng/China Mechanical Engineering, 2019, 30.2, pp. 244-251.

[6] Y. Chen, X. Du and L. Zhou, "Transformer defect correlation analysis based on Apriori algorithm," 2016 IEEE International Conference on High Voltage Engineering and Application (ICHVE). Chengdu, 2016, pp. 1-4.

[7] L. Shiyong et al., "Transformer Familial Defect Detection Method Based on SVM Improved by Apriori Algorithm—Based on Analysis of Dissolved Gases in Oil," 2019 2nd International Conference on Electrical Materials and Power Equipment (ICEMPE). Guangzhou, China, 2019, pp. 544-548.

[8] Y. Yang, Y. Geng, Y. Ju, X. Zhao and D. Yan, "Mining Association Rules from Multidimensional Transformer Defect Records," International Conference on Geo-Spatial Knowledge and Intelligence. Springer, Singapore, 2017, pp. 364-374.

[9] K. B. Sahay, A. Sonkar and A. Kumar, "Economic Load Dispatch Using Genetic Algorithm Optimization Technique," 2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE). Phuket, Thailand, 2018, pp. 1-5.

[10] M. Teixeira Magalhães, G. Brandão de Miranda, L. Brugiolo Gonçalves, L. Leão Oliveira Moreno, S. Sã Rosário Furtado Soares and L. Willer de Oliveira, "A Random Keys Genetic Algorithm Approach for Reconfiguration Problem in Distribution Power Networks," 2019 8th Brazilian Conference on Intelligent Systems (BRACIS). Salvador, Brazil, 2019, pp. 866-871.

[11] V. Suresh and P. Janik, "Optimal Power Flow in Microgrids using Genetic Algorithm," 2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC). Chennai, India, 2019, pp. 483-48.

[12] L. Zhendong et al., "Study on intelligent defect management strategy based on cloud mode," 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference. Chongqing, 2016, pp. 51-54.

[13] R. Agrawal, and R. Srikant. "Fast algorithms for mining association rules," very large data bases, 1998, pp. 580-592.

[14] H. S. Tsoukalas and R. E. Uhrig, "Fuzzy and Neural Approaches in Engineering," New York: Wiley, 1997.

Automatic Detection of Transmission Line on UAV Inspection Images with the Statistics Approach in the DCT Domain

Min Zhang
School of Mathmatics
Northwest University
Xi'an, China
dr.zhangmin@nwu.edu.cn

Abubakar Abubakar Khalid
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
abubakarakhalid@yahoo.com

Yifan Li
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
xgdlyf@163.com

Yu Chen*
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
chenyu@xjtu.edu.cn

Abstract—**Power transmission lines are significant aspect for uninterrupted electric energy supply. The aerial camera captured image processing methods use for inspecting the condition of this infrastructure involve algorithms that explicitly find line segments in the image and classify them according to width, length and angle. These methods are erroneous with extreme false detection outcome rate. It's obvious that image may contain a complex background with a line like structures. For this purpose, an alternative approach that visualizes the statistical model of local DCT coefficients is proposed to estimate the existence of power line. The proposed algorithm tackles the issue of power transmission line detection by exploiting low-level images characteristic to mimic human visual acuity and detect the existence of power line by mapping the extracted coefficient to Generalized Gaussian Distributions (GGDs) from the DCT coefficients. Then percentile pooling is applied on the fitted coefficient to obtain values representing the local and global distribution. Emphasis is laid on the orientation feature to capture the transmission line varying position and location in separate images. These features are feed as input to SVM for classification. The proposed algorithm was tested on the TEIAS visible light database containing 4000 undistorted images, and a local dataset containing 1039 distorted and undistorted images. Experimental results show the proposed method can detect power lines with reasonable performance.**

Keywords—Power Transmission Line Detection, Discrete Cosine Transform, Generalized Gaussian model, Support Vector Machine

I. INTRODUCTION

Image processing applications have found wide attentiveness in electric power automation field, such as printed circuit board, power transmission line inspection, signal transmission, etc. Power transmission networks play a crucial role in the realization of the enormous capacity and long-distance transmission of electricity [1]. In this regard, the field has been an involved research interest, with profound accessible literature and algorithms. Particularly, extensive automated inspection such as wheel robot, crawling robot and aerial imaging (Copter, Drone, SAR, etc.) systems were developed [2-5]. The conventional camera-based inspection extract edge features in image and map specified points assume to contain power lines by utilizing Hough transform [6]. Etemadi [7] suggests a line segmentation detector that can determine all segmentations of lines without depending on any parameter. Burns et al. [7] propose a detector merging pixel having the same orientation.

Correspondingly, Yetgin et al. have proposed a training-based approach that maps line segments in the image, then divide the lines into two groups, according to the length, angle and parallelism features of the lines [9].

Recently, researchers race toward developing rapid remote sensing techniques for UAVs platforms, considering the advantages such as obstacles avoidance, flexibility and reduce labor intensity [10-14]. Feng et al. [10] suggest the use of intuitive measurement through double-side filters and Hough transform with the parallel constraint for power line recognition of UAVs inspection. In 2014, Luo et al. developed a method that extracts combine structure, material, and color characteristics to express the properties of power lines with the aid of some corresponding image cues [11]. Li [12] used Pulse Coupled Neural Network (PCNN) filter to remove background noise from the images prior to the Hough transform and knowledge-based clustering is applied to summarize characteristics of the power line. Gioi et al. [13] developed an unsupervised line segments detector (LSD) with automatic error checking and utilizing pixel-wise processing to detect the slope angle along the line. Subsequently, Topal et al. make an improvement on the LSD for robust and fast detection and called it EDLines [14]. The major enhancement by the authors includes computational efficiency and line validation.

We proposed an algorithm which is capable of detecting if the image scene contains power transmission lines or not. The features utilized for this purpose have been selected from BLIINDS-II [15], a modified orientation feature by Yetgin et al. [16], and Support Vector Machine (SVM) classification [17]. Our model consists of four features including; Generalized Gaussian model shape parameter, Frequency variation, Energy ratio measure, and modified feature selection, pooled at three scales respectively. The four features are then integrated into an overall measure containing 24 column DCT modeled features representing thescene characteristics. Validations on the TEIAS visible light database [18] and our local dataset show the approach exhibits reasonable performance.

II. RELATED WORKS

In an extensive quest to establish a lasting solution to help low flying aircraft to detect power transmission cable, which is highlighted as the prime cause of accident. Ö.E. Yetgin and Ö.N. Gerek recommended a novel approach that elaborates

---

*Yu Chen is the corresponding author. (E-mail: chenyu@mail.xjtu.edu.cn)

on the DCT spectrum frames to interpret its frequency

signatures [16]. The authors approach the detection of power transmission line as scene classification, contrary to the swarm edge-line detection algorithms. In that aspect, the algorithm is simply trained to make a binary decision on images according to the existence of transmission cable using supervised classification.

The most essential image characteristics representing step is feature extraction. Recognizing the DCT spectrum successes in capturing both local and global feature in previous literature [15, 16]. The authors resolve to deploy the DCT to look for frequency signatures inside the image spectrum with full region search to monitor the inherent position and direction of the power cable. The algorithm subject test image to pre-processing, spatial to transform domain conversion, DCT coefficient selection and supervised classification. The classical selection and the proposed methods are as follows:

1) Classical Selection (CS): Knowledgeable about DCT frequency increase moving downward, the conventional approach selects features from the top-left part of the DCT as in Fig. 1 (a) [16], regarding most of the signal energy is contained in lower frequencies.

2) Reversed Selection (RS): Conversely, lines are expected to contain higher frequencies. Therefore, an approach worth trying is to select the DCT domain from the highest frequency region at similar varying domain sizes as shown in Fig. 1 (b) [16].

3) Patch-Based (PB): A natural alternative to the above two strategies is to select block regions over DCT samples in a tiling manner, as seen in Fig. 1 (c) [16]. This is logical due to the fact that the exact emphasis on DCT samples for images with power cables are arbitrary. Whenever the patch includes the top left corner, the DC component is omitted from the feature samples due to its irrelevance to power lines.

Finally, feature extraction and selection are implemented on MATLABTM [19], then transferred to Weka [20] software for classification purpose. The algorithm displays significantly better identification results, with lower false alarm rates.

III. THE PROPOSED POWER LINE DETECTION ALGORITHM

Synthesizing Saad et al. [15] and Yetgin et al. [16] philosophy, our approach consists of four steps: Pre-processing, Image coefficient extraction, DCT transformation and SVM classification. We present a description of these steps below:

A. Pre-processing

1) Image Resize: Considering that our aerial imaging gadget capture high-quality images saved in JPG format, with an average dimension of pixels [21]. Performing quality assessment task on these large images will consume significant time, which will render the algorithm inefficient. Considering the time efficiency, our images are resized to with negligible loss of quality in accordance with the work in [22, 23].

2) Color Space Transformation: Color is fundamentally a multi-dimensional phenomenon described by the perceptual
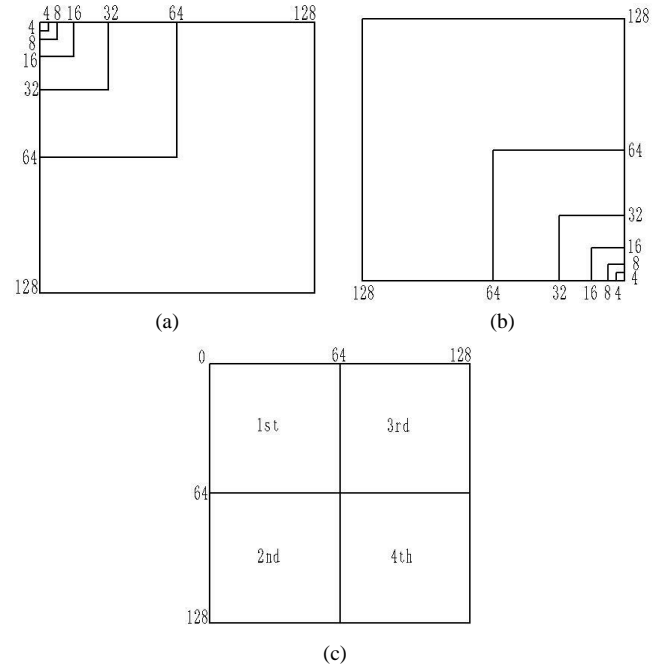
Fig. 1. DCT coefficient selection: (a) classical, (b) reversed selection and (c) patch based

attributes of luminance, chrominance and hue [22]. Most colors in the visible spectrum can be recreated by a combination of the RGB components. But in the RGB the structural information may be over/underestimated, due to highly correlated color components. We resolve to use the YCbCr color space. This choice is based on an experiment that luminance is more important than chrominance from a perceptual standpoint, the accuracy of the brightness information has far more impact on the image than that of the other two [23].

B. Discrete Cosine Transform

The benefit of representing an image in the DCT domain is that most of the visually significant information in the image is concentrated in just a few coefficients which represent frequencies instead of pixel values [15, 16]. By transforming the spatial coefficient to DCT spectrum, the DCT matrix is formed by splitting the image into equally sized $N \times N$ blocks with two-pixel overlap between neighboring blocks, which are represented as a sum of sinusoids with varying frequencies and magnitudes. Most of the image information is stored in the low-frequency region (top left corner of the matrix), while edge information is stored in the high-frequency region (bottom left corner of the matrix). This decomposition help in detecting image distortion by computing handcrafted statistical features like energy compaction, frequency variation, and then store this features in a matrix form. For each block, local parameter computation is performed and then add together for global estimation and percentile pooling to obtain two features.

C. DCT Signature and Patch Selection

Natural scene image coefficient in the DCT domain exhibits a predictable magnitude [16]. However, the cable position may arbitrarily appear in a different position. Yetgin et al. [16] proposed methods (Reversed Selection and block-

wise full region search of the entire DCT coefficient) serve as the suitable solution to detecting this variation in power line image. The method was exploited to counter the cable location and position uncertainty.

D. SVM Classification

Considering that a robust unbiased classifier with enhanced strategy of training-test approach is much desired for this task. SVM is chosen as the most suitable candidate, regarding its record achievement in previous literature [15, 24, 25]. The SVM requires a calibration procedure for robust performance. Such that no overlap between train and test content occurs. According to SVM [17] class label design architecture, our images are tag with two values "1 and 0". All training/test unadulterated natural scene images with power transmission line are labeled '1'. Likewise, unadulterated training images without or adulterated images (with or without) power transmission cable are labeled '0'. A prediction model is obtained after the completion of the train/test stages. This model is utilized to automatically check future transmission line image (with or without) corresponding decision, without repeating the train/test stages. The evaluation outcomes are presented in the experiment section.

IV. EXPERIMENTAL RESULTS

A. Power Transmission Line Datasets

To evaluate our method in a field batter with a meager testing platform, we resolve to the TEAS repository [18], containing 4000 visible light transmission images built by Yetgin et al. Obviously TEIAS database is designed specifically for testing the existence of power transmission cable in high-quality images scene, with no distorted version. These serve only a fraction of our objective. This issue motivates building a new local dataset which aims to significantly expand the diversity of images and has more image distortion categorical samples. Distortion incurred during field experiment and simulated distortion from original images. We generate images with four types of distortions.

1) Natural distortion

The main distortion observed in UAV transmission cable imaging include motion blur (caused by target/device fast movement 'or both'), blocking artifact (during compression & storage), low contrast and overexposure (Shutter speed). We collected 117 images affected by these distortions.

2) Simulated distortions

  a) Gaussian White Noise (53 images),

  b) Speckle (50 images),

  c) Gaussian Blur (51 images)

  d) Salt and Pepper (45 images).

B. Experiment

Our method was tested using linear SVM classifier in MATLAB environment, on a Lenovo computer running Windows 10 Pro 64-bit operating system, with Intel (R) Core (TM) i5-4210U CPU @ 1.70GHz 2.40GHz x64-based Processor, RAM: 4.00 GB. Multiple train–test sequences were run. In each experiment round the samples are partitioned into 70% and 30% test-train randomly. We do this to ensure that the reported results do not depend on features extracted from known content, which can artificially improve performance.

TEIAS dataset (4000 images) and a local dataset (1039 images) are decomposed into sets to train several SVM models to test our method efficiency. Some of the set only test image quality or existence of transmission line (or not) and then both image quality and transmission line in a scene. A set contains a minimum of 250 images, with their respective labels. The labels are used when training these models and testing for the prediction accuracy. The experiments are done in three phases: a model training phase, the prediction phase and accuracy check. Our goal is to train a model to mimic realistic behavior by first employing a collection of images of known quality, and then use them to estimate the quality of new images outside of the original collection.

The training process requires a sufficient number of samples for adequate learning (randomly selected to build our model) with known quality indices extracted using the approaches described above. Each training sample contains 24 column DCT modeled independent features representing the scene characteristics and respective label "1 or 0" corresponding to class belongingness. Usually, the parameters are randomly initialized. When predicting the class of arbitrary image samples, we consider the entire image and try to obtain DCT modeled features representing the scene characteristics to feed into the SVM as an input. If the set model is efficient, we expect the classification outcome to reflect initial training accuracy with the slightest decrease, due to noisy data points (outlier). On the other hand, if our outcome is far from the training. We consider the model as inefficient, we go for more training and possible parameters turning for a better model.

C. Results

Since SVM has both the image and its corresponding label, a loss function (Accuracy) can be defined. We used the Accuracy function to calculate the ground truth and the predicted total error of the output:

$$Accuracy = \frac{True\ Positive + True\ Negative}{All\ Samples}\% \qquad (9)$$

The concepts of true positive and true negative are illustrated in Fig. 2. The confusion matrix can display a better idea of what a supervised classification model is getting right and errors it is making. Significantly, it reveals not only the errors being made by a classifier but more importantly the types of errors that are being made. Table I display the outcome when tested on the two datasets.

TABLE I.   ACCURACY RESULT FROM THE TWO UTILIZED DATASETS

Image Source	Training Accuracy (%)	Prediction Accuracy (%)
TEIAS	82.6	73.95
Our Dataset	78.97	77.04

In Fig. 2 below, we show confusion matrix for random models, using the classification learner in MATLAB. The rows of the matrix represent the samples in a true class while the columns represent the samples predicted class (or interchange conversely).

In table II, we perform the experiments by training the based distortion and mixing both together. From the result, we can see higher accuracy is obtained when dealing with a single issue. Obviously, this expected considering each algorithm performance success in their original forms.
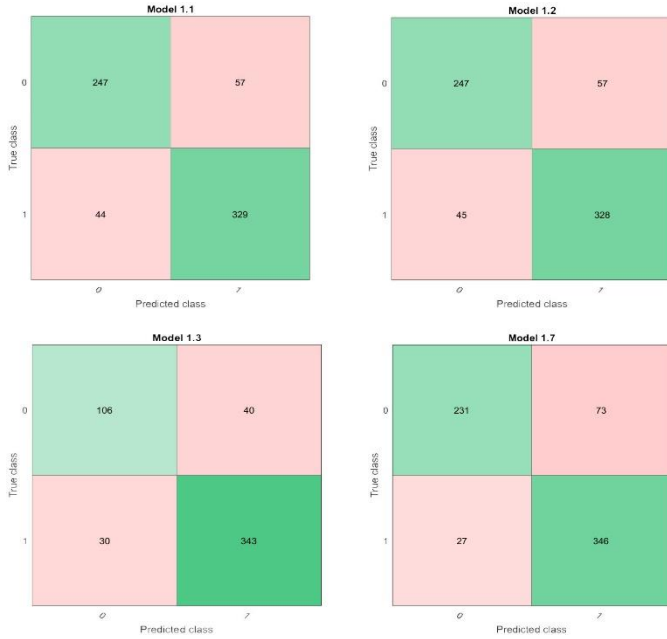

Fig. 2. Confusion matrix.

TABLE II. CATEGORICAL AND MIXED SVM ACCURACY

Image category	Training Accuracy (%)	Prediction Accuracy (%)
Cable Detection	82.79	76.47
Quality Evaluation	94.85	90.50
Mixed	77.89	75.14

To test our methods, we resolve to employ more classification method to observe the performance of our DCT features, Table III shows the results, utilizing 70/30% strategy of training-test isolation, using celebrated classifiers such as Linear Discriminant Analysis, Quadratic Discriminant Analysis, Coarse Tree and Fine Tree. Our approach (SVM) classifier display better performance. The computation time is negligible for all the classifier.

TABLE III. PROPOSED SVM METHOD COMPARED WITH OTHER SUPERVISED CLASSIFIERS

Algorithm	Prediction Accuracy (%)	Computational time(sec)
Linear Discriminant	83.9	11.455
Quadratic Discriminant	82.9	10.364
Linear SVM	85.2	11.949
Coarse Tree	84.9	9.9145
Fine Tree	85.1	10.221

Table IV shows the comparison of our method with Yetgin et al. which only is interested in the existence of transmission cable in a scene, while our method is extended to check both the image quality and existence of transmission cable in the scene. The two algorithms are tested on our dataset and TEIAS (VL) dataset. Of the three classifiers utilized by the authors (Naive Bayes, Random Frost and SVM), we choose SVM, considering our method of SVM classification. It important to note, that our method has a small feature representing the image. While the former use a full feature mapping.

TABLE IV. PROPOSED METHOD COMPARISON TO YETGIN ET AL., BASED ON POWER LINE DETECTION

Datasets	Yetgin	Proposed	Row-Column
TEIAS (VL)	87.23	84.71	128×128
local(unadulterated)	58.68	82.09	256×256

The performance result shows that Yetgin et al. approach is dependable on image size. If dimension data samples are to be used then the approach provides better performance. While it deteriorates for larger images, as observed when tested on our dataset. However beware that an exact comparison is not possible as though both approached are DCT based, the methods differ in objective and feature modeling.

V. CONCLUSION

This paper described an automatic detection algorithm for the power transmission line in images acquired with UAV. The algorithm consists of four features including: shape parameters of the Generalized Gaussian model; Frequency variation; Energy ratio measure and the feature selection, pooled at three scales respectively. The four features are then integrated into an overall measure containing 24 column DCT modeled features representing the scene characteristics. The algorithm was trained to mimic human prediction to distinguish unadulterated images with power line and images without power line or adulterated, on a combined dataset of 5,039 images. The experimental result shows the proposed method can detect power lines and evaluate image quality with an accuracy of 77%.

AKNOWLEDGEMENT

REFERENCES

[1] Dias R., Lima A., Portela C. and Aredes M., Extra Long-Distance Bulk Power Transmission, Transactions on Power Delivery, IEEE 2011, 26(3): 1440-1448.

[2] Lee H., Kim C., Park S., Lee S., Kim J., Kim S. and Lee T., Vision-based automatic real time inspection of power transmission line, International Symposium on Robotics, IEEE 2013, 44:1-3.

[3] Matikainen L., Lehtomaki M., Ahokas E., Hyyppa J., Karjalainen M., Jaakkola A., Kukko A. and Heinonen T., Remote sensing methods for power line corridor surveys, Journal of Photogrammetry and Remote Sensing, Elsevier 2016, 119: 10–31.

[4] Singh S. and Vishwakarma D. N., Intelligent techniques for fault diagnosis in transmission lines-An overview, International Conference

on Recent Developments in Control, Automation and Power Engineering, IEEE 2015, 280-285.

[5]   Martinez C., Sampedro C., Chauhan A. and Campoy P., Towards autonomous detection and tracking of electric towers for aerial power line inspection, International Conference on Unmanned Aircraft Systems, IEEE 2014: 284 – 295.

[6]   Duda R.O., Hart P.E., Use of Hough Transformation to detect lines and curves in pictures, Commun. of the ACM, vol.15, no.1, pp. 11–15, January 1972.

[7]   Etemadi A., Robust Segmentation of Edge Data, Image Processing and its Applications, IET 1992, 311-314.

[8]   Burns J.B., Hanson A.R. and Riseman E.M., Extracting Straight Lines, Transactions on Pattern Analysis and Machine Intelligence, IEEE 1986, (8)4: 425-455.

[9]   Yetgin O.E., Zekeriya Ş. and Gerek O. N., A comparison of line detection methods for power line avoidance in aircrafts, Electrical and Electronics Engineering, 2015, 9: 241-245.

[10]   Tian F., Wang Y. and Zhu L., Power Line Recognition and Tracking Method for UAVs Inspection, Proceedings on Information and Automation, IEEE 2015, 2136-2141.

[11]   Luo X., Zhang J., Cao X., Yan P. and Li X., Object-Aware Power Line Detection Using Color and Near-Infrared Images, Transactions on Aerospace and Electronic Systems, IEEE 2014, 50(2): 1374–1389.

[12]   Li Z., Liu Y., Hayward R., Zhang J. and Cai J., Knowledge-based Power Line Detection for UAV Surveillance and Inspection Systems, Machine Vision and Applications, Springer 2010, 21(5): 677–686.

[13]   Gioi R.V., Jakubowicz J., Morel J.M. and Ranball G., LSD: A Fast Line Segment Detector with a False Detection Control, Transaction on Pattern Analysis and Machine Intelligence, IEEE 2010, 32(4): 722-732 .

[14]   Topal C., and Akinlar C., EDLines: A real-time line segment detector with a false detection control, Pattern Recognition Letters, IEEE 2011, 32(13): 1633-1642.

[15]   Saad M., Bovik A. C., and Charrier C., Blind Image Quality assessment: A Natural Scene Statistics Approach in the DCT Domain [C], Transactions on Image Processing, IEEE 2012, 21(8): 3339–3352.

[16]   Yetgin O.E, Gerek, O.N., Automatic Recognition Of Scenes With Power Line Wires In Real Life Aerial Images Using DCT-Based Features, Digital Signal Processing. Elsevier 2017, pp. 1-18.

[17]   Cortes C., Vapnik V., Support-vector networks, Machine Learning, Kluwer Academic Publishers 1995, 20 (3):273–297.

[18]   Yetgin Ömer Emre; GEREK, Ömer Nezih (2017), Powerline Image Dataset (Infrared-IR and Visible Light-VL), Mendeley Data, v7, http://dx.doi.org/10.17632/n6wrv4ry6v.7.

[19]   MATLAB, http://www.mathworks.com, 2017 .

[20]   WEKA, Available: http://www.cs.waikato.ac.nz/ml/weka/, 2016.

[21]   Dà-Jiāng Innovations:https://www.dji.com.

[22]   Torralba A., How many pixels make an image?, Visual Neuroscience, Cambridge University Press 2009, 26 (01):123–131.

[23]   Meguro M., Takahashi C., and Koga T., Simple Color Conversion Method to Perceptible Images for Color Vision Deficiencies, Human Vision and Electronic Imaging XI, SPIE 2006, 6057: 432-442.

[24]   Z. Yazhong, W. Jinjian, X. Xuemei, L. Leida and S. Guangming, "Blind image quality assessment with improved natural scene statistics", Digital Signal Processing, vol. 57, pp. 56-65, Elsevier 2016.

[25]   P. Parashar and K Er. Harish, "Comparison of various image classification methods", International Journal of Advances in Science and Technology, vol. 2, no. 1, 2014 .

Risk Grading of Distribution Network Equipment Group Based on Fuzzy Clustering Factor Analysis

Hongbin Wang
Guangdong Power Grid Co., Ltd
Guangzhou power supply
GuangZhou, China
13503095341@139.com

Le Luan
Guangdong Power Grid Co., Ltd
Guangzhou power supply
GuangZhou, China
149529958@qq.com

Kai Zhou
Guangdong Power Grid Co., Ltd
Guangzhou power supply
GuangZhou, China
285358753@qq.com

Xueyan Chai*
Shaanxi Key Laboratory
of Smart Grid
the State Key Laboratory of
Electrical Insulation and
Power Equipment
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
1803563707@qq.com

Xue Li
Shaanxi Key Laboratory
of Smart Grid
the State Key Laboratory of
Electrical Insulation and
Power Equipment
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
lx6fd7663@163.com

Minghui Xie
Shaanxi Key Laboratory
of Smart Grid
the State Key Laboratory of
Electrical Insulation and
Power Equipment
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
619425976@qq.com

Abstract—The power system, however, has shifted from a regulated system to a competitive, uncertain market environment. Various risks threaten the safe and stable operation of the distribution network at any time. At present, the research objects of risk assessment are mainly single equipment and overall system. However, it is difficult for these two research results to directly reflect the risk situation, and give suggestions for operation and maintenance improvement based on the results. At the same time, traditional risk assessment methods rely too much on expert advice. The actual evaluation efficiency is low and the application is difficult to promote. This paper proposes a method of risk grading for distribution network equipment group based on fuzzy clustering factor analysis. It can flexibly allocate network equipment groups and use equipment groups as research units. Based on multi-factor and multi-level distribution network equipment risk information, we refine the risk assessment index information by refining public factors. This method can not only optimize the risk assessment index information of distribution network equipment, but also improve the accuracy of assessment. And from the perspective of the equipment group, it is more convenient for the risk assessment conclusion to be applied to the distribution network, which has great practical significance. Some feeder data of Guangzhou are used to verify this method. The experimental results indicated that this method can be more easily applied to the actual distribution network operating environment.

Keywords—risk assessment, distribution network equipment, factor analysis, state evaluation Introduction

As the end of the power system, the distribution network is directly connected to users. It is one of the key links of the power grid. The network coverage is wide and the topology is complex. It is closely related to the normal development of industrial production and people's lives. As the load types and the system components increase, users' power requirements become more demanding, and it is more difficult to manage the distribution network. The possibility of random failures is becoming more and more prominent. The statistical results show that 80% of power grid failures are caused by distribution network failures. Once the power failure occurs, the losses brought to people's daily life and social

development become more serious than before. In order to ensure the safe and stable operation of the distribution network, it is necessary to conduct a risk assessment of its equipment in terms of the distribution network operation risk. To establish a reasonable risk assessment and grading model, we would analyze various factors to achieve considerable and controllable risks.

In the 1960s, foreign countries first began research on the reliability and risk of distribution networks. The safety and economy of the distribution network are unified to comprehensively reflect the risks, and the relevant research is relatively mature [1]. Our country's power industry once only paid attention to the safety and stability of the transmission system to ensure the sufficiency of power supply. The research about safety and stability of the distribution network is usually ignored. It results a relatively weak grid structure, and power outages and restrictions both have a high frequency. There are many aspects about the traditional power system risk grading, such as determining the component outage model, selecting the system state, calculating their probabilities, evaluating the consequences of the selected state, and calculating risk indicators. Reference [2] from the perspective of risk assessment theory, it determines the component outage model and load curve model firstly, followed by enumeration sampling method to calculate the probability and impact of various states of the distribution network system. Finally, it calculates the load loss probability and power shortage. Reference [3] applies the depth-first search method to fault enumeration and consequence analysis, which greatly reduces the workload and calculation time. The traditional risk grading method is limited by few consideration factors, only considering from the two aspects of component level and system level, which reduces the practicability of the distribution network and has gradually been unsuitable for the development of smart grid.

In order to improve the shortcomings of the above existing methods, the purpose of this paper is to provide a method for risk grading of distribution network equipment group based on public factor fuzzy clustering. Based on multi-factor and multi-level distribution network equipment risk information, it can scientifically and effectively simplify the original

---

* Xueyan Chai is the corresponding author. (e-mail: 1803563707@qq.com).

information, ensure linear independence. It accomplishes an overall objective and reasonable assessment of distribution network equipment risk. This method can not only optimize the risk assessment index information of the distribution network equipment, but also improve the accuracy of the evaluation. From the perspective of the equipment group, it is more convenient for the guidance of the risk assessment conclusion to be applied to the distribution network.

This paper is organized as follows: In Section 2, we provide research background on risk analysis and the factor analysis. In Section 3, the proposed method is demonstrated step by step. In Section 4, the experiment datasets are firstly demonstrated and then experiment results are analyzed to validate our proposed method available. Section 5 concludes this paper.

I. EASE OF USE

A. Maintaining the Integrity of the Specification

Risk, representing system health, is a combination of the possibility of uncertain events and their consequences. Broadly speaking, risk can be defined as $R = (H, P, C)$ [4]. R is the risk. H is the hazard. P is the probability of the occurrence of the hazard, and C is the cost of the loss caused by the hazard. In actual work, the product of the probability of a dangerous event and the resulting loss is often used as the expected value of risk to describe the severity of the loss caused by the dangerous event to society. This expected value is the risk in the narrow sense, which can be described as $R=f(P, C)$ [5].

The essence of risk in the distribution network lies in the probability of its behavior. Most of factors are difficult to predict accurately, such as the effects of random equipment failure, load uncertainty, external and human factors. These factors may cause large-scale power outages in the system that are local or even entire. From the perspective of power companies, the risks of distribution networks can be divided into external risks and internal risks [6]. External risks are mainly reflected in the environmental risks of power grid operation. There are many risks in the distribution network operation and production process including natural, social, technical, environmental and personal aspects, such as natural disasters, outdated equipment, and weak distribution grid structure [7]. External risks are also reflected in the reform of electricity marketization.

Power system reform will bring difficulties in power dispatch management and coordination. The resulting grid security risks and competition will lead to a decline in grid security and stability. The internal risks of distribution network operation mainly include operation risk, management practice risk and personnel quality risk. Therefore, our main research is the environmental risk of distribution network operation. According to practical experience, equipment failure is the main cause of power outages in distribution networks [8]. Therefore, this paper mainly focuses on the risks caused by equipment failure in the distribution network operating environment.

The operation of the distribution network needs to take into account both safety and economy. The traditional deterministic safety assessment method is based on the "most serious accident decision criteria", which means the system must ensure safe operation under extreme operating conditions and the most serious faults [9]. Therefore the research results are often too conservative to meet the economic requirements of system operation.

Distribution network risk assessment comprehensively considers the probability of accidents and the consequences of accidents. It controls the operation of the system by considering the risks of the system, so that the safety and economy are controlled within an acceptable range. Since the distribution network risk assessment theory was put forward, domestic and foreign scholars have done a lot of research on it. Research hotspots mainly focus on the basic concepts and importance of risk assessment [10], construction and improvement of risk assessment index system and risk-based control decision [11].

With the continuous deepening of engineering applications, some defects of the risk indicator system gradually appeared. They are mainly reflected in insufficient assessment details and insufficient risk information. Therefore, the dispatcher can not understand details of the origins for the risks and the consequences of the risks through the risk indicator value. They cannot understand the proportion of the risk value under different predicted accidents in the comprehensive risk results, and it is difficult to qualitatively estimate the criticality of the risk. These problems hinder the further promotion and application of the risk index system in engineering practice.

With the deepening of research, research scholars hope to determine the criticality of risks by grading risk indicators. Reference [12] comprehensively obtains the system safety level based on the system low voltage risk index, system overload risk index and system voltage instability risk index. Reference [13] uses fuzzy comprehensive assessment method to assess and grade the risk of accident chain. References [14] establish a risk rating system for transmission lines. However, due to the few types of risks and the factors that affect the risks considered, these achievements are not mature enough in terms of system application. Therefore, this paper integrates the existing risk grading methods, based on the xx criterion, comprehensively considers the risk impact factors from a practical perspective. And we take the equipment group as the research unit, so that the risk grading results can be more effectively applied to guide the distribution network operation.

B. The Factor Analysis Method

The factor analysis method is a multivariate statistical analysis method that summarizes some variables with intricate relationships into a few unrelated new comprehensive factors [15]. The basic idea is to group variables according to the size of the correlation, so that the correlation between variables in the same group is higher, and the correlation between variables in different groups is lower. Each set of indicator variables can be represented by an unmeasurable comprehensive variable, which is called a common factor. In this way, a few representative public factors can be used to summarize the information provided by the multivariate to find out the main factors that affect the observed data. Assume that the evaluation population has k evaluation indicators and n observation units. In the multi-index comprehensive evaluation method, the traditional method often has a certain subjective arbitrariness in setting the weight [16]. Applying the factor analysis method to comprehensive evaluation can overcome the shortcomings of weight setting, high correlation of index information and information redundancy. It also makes the comprehensive evaluation result unique and objective. Therefore, in the process of exploring the risk

grading of the distribution network, we use factor analysis to balance the multi-risk indicators under comprehensive consideration, refine the risk indicator system, and reduce the amount of useless calculations. The mathematical model of factor analysis is to express n observation units as a linear weighted sum of p common factors and a special factor, where p<k. And it can be described as:

$$\chi_i = \alpha_{i1}F_1 + \alpha_{i2}F_2 + \cdots + \alpha_{ip}F_p + \varepsilon_i \left(i = 1,2,\cdots,n\right) \quad (1)$$

where F_1, F_2, ..., F_P are common factors. They are usually independent of each other. ε_i is a unique factor for each corresponding variable χ_i, called a special factor. We usually assume that $\varepsilon_i \sim N(0, (\sigma_i)^2)$ The coefficient α_{ij} is the coefficient of the *i-th* variable for the *j-th* common factor, called factor load. It reveals the relative importance of the *i-th* variable for the *j-th* common factor.

II. Algorithm

We divide the equipment group according to the different attributes of the distribution network equipment. For example, feeders, villages in cities, etc., which can be determined according to actual needs. Assume that there are n risk assessment objects for the distribution network equipment group and u risk assessment indicators, and the sample data set is *{aij|i=1,2,...,n;j=1,2,...,u}*. *aij* indicates the value of the *i-th* evaluation object corresponding to the *j-th* evaluation index.

In order to solve the problem of distance calculation due to different dimensions or different units, the original sample data is standardized [17]. The standardized formula is as follows:

$$\tilde{a}_{ij} = \frac{a_{ij} - \mu_j}{s_j} \quad (2)$$

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} a_{ij} \quad (3)$$

$$s_j = \frac{1}{n-1}\sum_{i=1}^{n}\left(a_{ij} - \mu_j\right)^2 \quad (4)$$

where $\tilde{a}_{ij}$ is the standardized data; μ_j is the average value of the *j-th* evaluation indicator variable; s_j is the standard deviation of the *j-th* evaluation indicator variable. Calculate the correlation coefficient matrix R of risk assessment indicators. The calculation formula is as follows:

$$R = \left(r_{ij}\right)_{u\times u} \quad (5)$$

$$r_{ij} = \frac{\sum_{k=1}^{n}\tilde{a}_{kt}\cdot\tilde{a}_{kj}}{n-1}\left(t, j = 1,2,...,m\right) \quad (6)$$

Where r_{tj} is the correlation coefficient between the *t-th* index and the *j-th* index, and $r_{tj} = r_{jt}$, $r_{jj} = 1$.To calculate the elementary load matrix, we need to solve the characteristic equation $\lambda I - R = 0$ to get the characteristic value λ_t *(t=1,2,...,u)* and $\lambda_1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_m \geqslant 0$. Then we find the eigenvector $\bar{x}_t$ *(t=1,2,...,u)* of the eigenvalue λ_t. The primary load matrix is obtained as follows:

$$\Lambda = \left[\sqrt{\lambda_1}x_1,...,\sqrt{\lambda_u}x_u\right] \quad (7)$$

Now we can seek the common factor of the risk index. Generally, the selection of public factors has the following three criteria: (1) Use the characteristic value of the principal component as the standard to select the common factor. We

can select the principal component with the characteristic value greater than or close to 1 as the common factor, and discard the other principal components with the characteristic value far less than 1; (2) The common factor is selected based on the standard cumulative variance contribution rate of the principal component. The cumulative variance contribution rate reflects how much the principal component retains the original amount of information. Generally speaking, the problem can be well explained by the cumulative contribution rate of the principal component reaching 85% or more. (3) The common factor is selected according to the specific problem or its professional theory. We take the second method. Among the u risk assessment indicators, we select the first, second, ..., *k-th* indicators corresponding to the cumulative contribution rate as public factors, which satisfies the formula: $\Sigma^k{}_{i=1}/\Sigma^u{}_{i=1} \geqslant 85\%$.Using the maximum variance rotation method to rotate the k common factor load matrix, we get Λ'. And the sum of the relative errors of the squares of all the elements of column k, which is $\Lambda' = \Lambda_1' + \Lambda_2' + ... + \Lambda_k'$, reaches the maximum. The calculation formula is as follows:

$$\Lambda' = \Lambda^{(k)}T \quad (8)$$

$$V_j = \frac{1}{u}\sum_{p=1}^{u}\left(d_{pj}^2 = \bar{d}_j\right)^2 \quad (9)$$

where T is an orthogonal matrix. $\Lambda^{(k)}$ is the first *k* columns of Λ. V_j is the relative variance of the square of the element in column j of the common factor load matrix Λ'. $d_{pj}^2 = \Lambda'_{pj}^2 / h_p^2$, $\bar{d}_j = \frac{1}{u}\sum_{p=1}^{u}d_{pj}^2$, $h_p^2 = \sum_{j=1}^{k}\Lambda_{pj}'^2$, d_{pj} and $\bar{d}_j$ are intermediate variables. h_p^2 is the common variance, reflecting the contribution of common factors to the original indicators.Using the Thomson regression method to calculate the score of *k* common factors, the calculation formula is as follows:

$$\hat{F} = AR^{-1}\Lambda' = \left(\hat{F}_{ij}\right)_{n\times k} \quad (10)$$

where *A* is the normalized matrix of the original data.$\hat{F}_{ij}$ is the estimated value of the *i-th* sample data for the *j-th* factor. After obtaining the common factor information, further classify the device groups through clustering methods to discover the internal connections among multiple device group objects. This is a fuzzy clustering problem with unknown number of clusters [18]. We choose fuzzy C-means clustering analysis method to solve the fuzzy clustering problem. Set fuzzy weighted index is *m (m>1)*, and the number of cluster centers is *e (2 ≤ e ≤ n)*. The iteration stop threshold is *x*, and the specified number of iterations is *b*. Initialize the membership matrix *P* with a random number between *0* and *1*, and they meet the constraints shown below:

$$\sum_{i=1}^{e} p_{ij} = 1, \forall j = 1,...,n \quad (11)$$

Calculate *e* cluster centers using common factor score $\hat{F}$, and *ei=1,...,e*. The cluster center can be calculated as follows:

$$e_i = \frac{\sum_{j=1}^{n} p_{ij}^m f_j}{\sum_{j=1}^{n} p_{ij}^m} \quad (12)$$

Determine if the value of the value function is less than the threshold or the amount of change from the previous value function value is less than the threshold, then stop iterating. The value function calculation formula is as follows:

$$J(P, e_1, ..., e_e) = \sum_{i=1}^{e} \sum_{j}^{n} p_{ij}^m d_{ij}^2 \qquad (13)$$

Where $d_{ij} = \| e_j - f_j \|$ is the Euclidean distance between the *i-th* cluster center and the *j-th* data point. Use the following formula to update the membership matrix P.

$$p_{ij} = \frac{1}{\sum_{c=1}^{e} \left(\frac{d_{ij}}{d_{cj}} \right)^{\frac{2}{(m-1)}}} \qquad (14)$$

Continuously calculate the clustering center and perform iterative calculation until the constraint conditions are met. Assess the risk level of different distribution network equipment groups according to the membership matrix P. The risk assessment of the distribution network equipment group is realized, thus we complete a risk grading process of the distribution network equipment group based on the fuzzy clustering of common factors.

III. EXPERIMENT

We select feeder lines as equipment groups to testify our method and get their risk levels. We choose the evaluation date to be 2019-10-11. We extract fields related to risk assessment indicators from the data source system, such as the weather, loss of power, important users, construction state and so on. Part of the original field data is shown in the following table:

TABLE I. ORIGINAL DATA OF SOME EQUIPMENT GROUPS

Location	Downtime	Power Restoration Time for users	Climatic Conditions	Loss of Load(kWh)
BaoGang F2	2018/9/1 11:38	2018/9/1 12:42	Rainstorm	1824.03
BaoGang F11	2018/9/1 11:39	2018/9/1 13:13	Rainstorm	5784.04
ChiGang F30	2018/9/3 15:11	2018/9/4 18:10	Thunder -shower	15035.88
RuiBao F16	2018/9/16 1:18	2018/9/16 2:17	Rainstorm	4220.54
NanJi F9	2018/9/23 15:23	2018/9/23 17:39	Thunder -shower	6348.95
XinZao F16	2018/9/12 9:14	2018/9/12 20:11	Sunny	12214.28
LiCun F40	2018/9/15 8:47	2018/9/15 10:46	Sunny	9631.12

The data has missing values, duplicate and incorrect values. We need to clean the data, calculate the risk assessment indicators. There are also some text data, which needs to be converted into numeric data. Finally, we get the risk assessment index data set, which are more than 10,000 data. The entire process diagram is shown below:
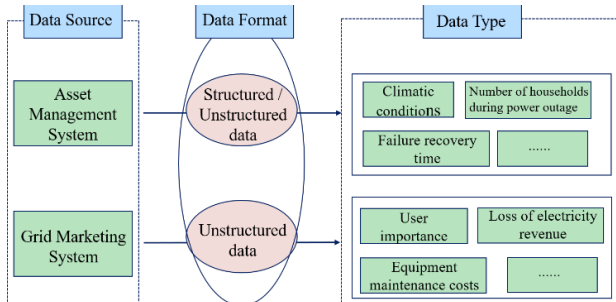

Fig. 1. How to extract indicator data.

The sample data set is expressed as $\{a_{ij} | i=1,2,...,10774; j=1,2,...,10\}$, a_{ij} represents the value of the *i-th* evaluation object corresponding to the *j-th* evaluation index. Calculate the correlation coefficient matrix R of risk assessment indicators. Solve the characteristic equation $\lambda I - R = 0$. We can get the eigenvalues. Find the eigenvector $\bar{x}_t$ $(t=1,2,...,10)$ corresponding to λ_t. The elementary load matrix is $\Lambda = [\sqrt{\lambda_1} x_1, ..., \sqrt{\lambda_u} x_u]$. Choose 3 indicators that satisfy $\sum_{i=1}^{k} / \sum_{i=1}^{u} \geqslant 85\%$ as common factors. Use the maximum variance rotation method to rotate the three common factor load matrices to obtain Λ'. And the sum of the relative errors of the squares of all three columns of Λ'. reaches the maximum. Calculate the scores of 3 public factors using Thomson regression. Set the fuzzy weighted index m to 2, the number of clustering centers to 3, the iteration stop threshold to ε. Specify the number of iterations are 50, and initialize the membership matrix P with a random number between 0 and 1. And it satisfies the constraints condition: $\sum_{i=1}^{e} p_{ij} = 1, \forall j = 1, ..., n$. Use common factor score $\hat{F}$ to calculate 3 cluster centers. Continuously update the iterative membership matrix until the value function value is less than the threshold, or the amount of change from the previous value function value is less than the threshold. Assess the risk level of different distribution network equipment groups according to the membership matrix P. The risk grading results of some equipment groups are shown in Table 2 below:

TABLE II. RISK GRADING RESULTS OF SOME EQUIPMENT GROUPS

Location	Equipment Group	v₁	v₂	v₃	Risk Level
HaiZhu	BaoGang F2	0.012	0.119	0.869	High
HaiZhu	ChiGang F30	0.023	0.037	0.940	High
HaiZhu	RuiBao F16	0.009	0.059	0.932	High
HaiZhu	NanJi F9	0.000	0.001	0.999	High
PanYu	XinZao F16	0.009	0.020	0.971	High
PanYu	LiCun F40	0.002	0.005	0.993	High
LiWan	LiuHua F11	0.992	0.007	0.002	Medium
HaiZhu	DaTang F21	0.973	0.021	0.006	Medium
BaiYun	NanDe F1	0.972	0.023	0.005	Medium
PanYu	ShiQiao F35	0.993	0.006	0.001	Medium
PanYu	DeXing F10	0.940	0.054	0.007	Medium
HuaDu	FuRong F23	0.975	0.023	0.002	Medium
ZengCheng	YongHe F19	0.989	0.008	0.003	Medium
HaiZhu	LiYao F5	0.016	0.976	0.008	Low
BaiYun	XiaMao F8	0.022	0.937	0.040	Low
BaiYun	ChaTou F1	0.032	0.932	0.036	Low
BaiYun	XiaoPing F23	0.016	0.976	0.008	Low
BaiYun	JiXian F3	0.003	0.996	0.001	Low
NanSha	HuDiezhou F16	0.000	0.999	0.000	Low

ZengCheng	XinTang F5	0.004	0.993	0.003	Low
CongHua	LongXing F13	0.010	0.986	0.004	Low
HuangPu	GanZhu F16	0.005	0.994	0.002	Low
TianHe	LongDong F4	0.038	0.923	0.039	Low

We get the membership degree matrix. According to the maximum membership degree principle, we can determine which risk level the feeder belongs to every day.This result shows the risk grading results for each equipment group (the equipment group is the feeder here). These results are based on the risk index system, which scientifically reflects the degree of risk. We compare the results with the results obtained by the existing grid risk assessment methods, and only 0.27\% of the results are inconsistent. At the same time, the calculation speed of the proposed method is 4 times that of the existing grid risk assessment method. Obviously, the method proposed in the article can greatly reduce the computational workload. It is no longer necessary to score and assess risks on a device-by-device basis, and there is no need to manually set different weights according to various devices. This greatly improves the efficiency and accuracy of risk assessment. Subsequent operation work can be carried out by the equipment group as a unit for maintenance, which greatly improves its practical applicability.

IV. CONCLUSION

The use of risk grading method has shown its practicality. This paper proposes a method of risk grading for distribution network equipment group based on fuzzy clustering factor analysis. It greatly streamlines risk assessment indicator information. An experiment is performed on feeder data to verify its feasibility. It solves the conflict between the redundancy of indicator information, the complexity of calculation, the lack of indicator information in the risk assessment of traditional distribution network equipment and the low reliability of the assessment results. An improved algorithm based on public factor fuzzy clustering is adopted to reasonably and effectively consider the correlation between risk assessment indicators, which improves the reliability of the algorithm. Compared with the traditional distribution network equipment risk assessment method, the method proposed in this paper can well integrate multi-level and multi-factor risk assessment indicators and improve the assessment efficiency and accuracy. It cuts in from the perspective of distribution network equipment group, which improves the risk assessment benefit of distribution network equipment and increases its practical application significance.

However, the risk level we have assessed is actually a quantification of historical risk based on the assessment date. If we want to learn about more intelligent distribution network situation awareness, we should study risk warning. Make preparations and adjustments in advance to maximize the use value of resources. We can reduce the probability of risk occurrence from the source and the loss after the risk occurs, and improve the reliability of users' electricity.

ACKNOWLEDGEMENT

REFERENCES

[1] L. Guo, C. Guo, Y. Cao, X. Jiang, and K. Sheng, "Power transmission risk assessment considering breaker online status," Automation of Electric Power Systems, vol. 36, no. 16, pp. 20–24+30, 2012.

[2] W. Qiu, J. Zhang, N. Liu, X. Zhu, and L. Liu, "Multi-objective optimal generation dispatch with consideration of operation risk," Proceedings of the Csee, vol. 32, no. 22, pp. 64–72, 2012.

[3] Yong Jiang, J. D. McCalley, and T. Van Voorhis, "Risk-based resource optimization for transmission system maintenance," IEEE Transactions on Power Systems, vol. 21, no. 3, pp. 1191–1200, 2006.

[4] D. Wang and Y. Yu, "Dynamic security risk assessment and optimization of power transmission system," in Proceedings of the 10th International Conference on Probablistic Methods Applied to Power Systems, 2008, pp. 1–6.

[5] H. Jia, W. Qi, Z. Liu, B. Wang, Y. Zeng, and T. Xu, "Hierarchical risk assessment of transmission system considering the influence of active distribution network," IEEE Transactions on Power Systems, vol. 30, no. 2, pp. 1084–1093, 2015.

[6] J. D. McCalley, A. A. Fouad, V. Vittal, A. A. Irizarry-Rivera, B. L.Agrawal, and R. G. Farmer, "A risk-based security index for determining operating limits in stability-limited electric power systems," IEEE Transactions on Power Systems, vol. 12, no. 3, pp. 1210–1219, 1997.

[7] Peng Wang and R. Billinton, "Reliability cost/worth assessment of distribution systems incorporating time-varying weather conditions and restoration resources," IEEE Transactions on Power Delivery, vol. 17, no. 1, pp. 260–265, 2002.

[8] A. A. Fouad, Qin Zhou, and V. Vittal, "System vulnerability as a concept to assess power system dynamic security," IEEE Transactions on Power Systems, vol. 9, no. 2, pp. 1009–1015, 1994.

[9] F. Xiao and J. D. McCalley, "Power system risk assessment and control in a multiobjective framework," IEEE Transactions on Power Systems, vol. 24, no. 1, pp. 78–85, 2009.

[10] J. D. McCalley, V. Vittal, and N. Abi-Samra, "An overview of risk based security assessment," in 1999 IEEE Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.99CH36364), vol. 1, July 1999, pp. 173–178 vol.1.

[11] D. Kirschen and D. Jayaweera, "Comparison of risk-based and deterministic security assessments," Generation Transmission Distribution Iet, vol. 1, no. 4, p. 527, 2007.

[12] Ming Ni, J. D. McCalley, V. Vittal, and T. Tayyib, "Online risk-based security assessment," IEEE Transactions on Power Systems, vol. 18, no. 1, pp. 258–265, 2003.

[13] H. J. Shi, F. Ge, M. Ding, R. L. Zhang, and H. Lin, "Research on online assessment of transmission network operation risk," Power System Technology, vol. 29, no. 6, pp. 43–48, 2005.

[14] W. H. Chen, Q. Y. Jiang, Y. J. Cao, and Z. X. Han, "Risk assessment of voltage collapse in power system," Power System Technology, vol. 29, no. 19, pp. 6–11, 2005.

[15] C. Reimann, P. Filzmoser, and R. G. Garrett, "Factor analysis applied to regional geochemical data: Problems and possibilities," Applied Geochemistry, vol. 17, no. 3, pp. 185–206, 2002.

[16] W. Li, Probabilistic Transient Stability Assessment, 2005, pp. 263–280.

[17] R. F. Ghajar and R. Billinton, "Economic costs of power interruptions: a consistent model and methodology," International Journal of Electrical Power Energy Systems, vol. 28, no. 1, pp. p.29–35, 2006.

[18] F. Xiao and J. D. McCalley, "Power system risk assessment and control in a multiobjective framework," IEEE Transactions on Power Systems, vol. 24, no. 1, pp. p.78–85, 2009.

Temperature Online Monitoring System for Aerospace Manufacturing Process Based on Gradient Boosting Decision Tree (GBDT) Algorithm

Liliang Wang
*School of Instrumentation and
Optoelectronic Engineering
Beihang University*
Beijing, China
wangliliang@buaa.edu.cn

Jiaqi Qu
*School of Instrumentation and
Optoelectronic Engineering
Beihang University*
Beijing, China
qujiaqi@buaa.edu.cn

Zheng Qian*
*School of Instrumentation and
Optoelectronic Engineering
Beihang University*
Beijing, China
qianzheng@buaa.edu.cn

Abstract—**The temperature measurement is important in aerospace manufacturing, with the increasing demand for online temperature monitoring. However, it is difficult to obtain high accuracy in traditional temperature detection methods using look-up tables and hardware compensation methods. This paper proposes an online temperature monitoring method for thermocouples based on the gradient boosting decision tree (GBDT) algorithm which enhanced the accuracy of cold end compensation and reduced nonlinear error of thermocouple during online temperature monitoring. Through experimental measurement, it can be seen that the accuracy of thermocouple online temperature monitoring improved by the gradient boosting decision tree algorithm provides a new method for online temperature monitoring in the aerospace manufacturing process.**

Keywords—gradient boosting decision tree，aerospace manufacturing，temperature，Thermocouple

I. INTRODUCTION

In aerospace, temperature measurement is very important research. For example, in the thermal vacuum test of high-power communication satellites, the temperature accuracy is required to reach ± 0.5 ℃ [1], the electrostatic discharge of GEO orbital solar cell array will decrease with the increase of temperature discharge frequency [2], the dual-line array stereo mapping camera of the Gaofen-7 satellite is very sensitive to temperature, and temperature changes will directly affect the image quality of the camera [3], and various instruments and equipment in the spacecraft need to work at a suitable temperature level. Therefore, accurate measurement of temperature is very necessary. Thermocouple is one of the most commonly used temperature detection components in industry [4]. In the use of thermocouple technology to measure temperature, there are many references [5-6]. The working mechanism of the thermocouple is the electrothermal effect of the conductor, including the Peltier effect and the Thomson effect, which are divided into contact thermoelectric potential and temperature difference thermoelectric potential. The temperature to be measured is calculated by measuring the thermoelectric potential [7]. Thermocouple has the advantages of simple structure, good stability and high sensitivity [8]. And traditional thermocouples can pass 0℃ constant temperature method, correction method, compensation bridge method, extension lead Methods such as hot electrode for error compensation [7]. It is also possible to use methods such as the analysis of the thermocouple signal acquisition circuit and the non-linear processing of the thermocouple signal to improve the temperature measurement

accuracy [9]. However, due to the accumulated errors in the experimental circuit, cold end compensation, and nonlinear correction, the measurement accuracy is insufficient. At present, machine learning has many research results in the field of thermocouple temperature measurement and correction [10]. For example, Tao Xiufeng uses a support vector machine model to eliminate the nonlinear error of thermocouples, and the party election uses neural networks and other models to perform nonlinear corrections on thermocouples. Considering that the neural network is not the global optimal solution, there will be a local optimal solution [11], which will make the temperature measurement results less accurate. This paper is based on the gradient boosting algorithm in machine learning to correct the temperature. The gradient boosting decision tree (GBDT) algorithm and its optimized algorithm Xgboost have the advantages of high prediction accuracy, fast training speed and strong robustness [12], so use gradient boosted decision trees and Xgboost algorithm fit the test data and correct the temperature.

II. THE PRINCIPLE OF GRADIENT BOOSTING DECISION TREE ALGORITHM

Decision tree is an algorithm based on a tree graph, which obtains leaf nodes through continuous splitting, and the data of leaf nodes can be manipulated in different ways, so the decision tree algorithm has classification and regression functions [13]. Decision tree iteration methods include ID3, C4.5 and CART algorithms. ID3 algorithm is based on information gain [14], C4.5 algorithm is based on information gain ratio split [15], and CART tree is based on Gini coefficient [16], which can be used Come classification and regression.

Boosting algorithm is a kind of integrated learning, which can boost the performance of weak learners. It uses training data to acquire a series of weak learners, and then combines these weak learners in some way. The model effect is significantly better than the initial weak learner. The algorithm steps [17] are:

Initialize the weight distribution of the training data:

$$D_1 = (W_{1,1},...,W_{1,N}), W_{1i} = \frac{1}{N} \qquad (1)$$

Train the base learner using data with a weight of D_m：

$$G_m(x): X \rightarrow \{-1,1\} \qquad (2)$$

Calculate the classification error rate of $G_m(x)$:

$$e_m = \sum_{i=1}^{N} P(G_m(x_i) = y_i) \quad (3)$$

Calculate the coefficient of Gm $G_m(x)$:

$$\alpha_m = \frac{1}{2} \log \frac{1-e_m}{e_m} \quad (4)$$

Update the weight distribution of the training set:

$$D_{m+1} = (W_{m+1,1}, ..., W_{m+1,N}) \quad (5)$$

Construct a linear combination of basic classifiers:

$$f(x) = \sum_{m=1}^{N} \alpha_m G_m(x) \quad (6)$$

Finally get a strong learner model:

$$G(x) = sign(f(x)) \quad (7)$$

The boosting algorithm with CART tree as the initial weak learner is called boosting tree. The linear combination model of the decision tree can be used to interpret the boosted tree algorithm:

$$f_M(x) = \sum_{m=1}^{M} T(x; \theta_m) \quad (8)$$

The initial tree is $f_0(x) = 0$, and the model of the m-th step is:

$$f_M(x) = f_{m-1}(x) + T(x; \theta_m) \quad (9)$$

Minimize experience risk:

$$\min \frac{1}{N} \sum_{i=1}^{N} L(y_i, f(x_i)) \quad (10)$$

The forward step algorithm is as follows:

$$f_0(x) = 0 \quad (11)$$

$$f_m(x) = f_{m-1}(x) + T(x; \theta_m) \quad (12)$$

$$f_M(x) = \sum_{m=1}^{M} T(x; \theta_m) \quad (13)$$

Step m, get $f_{m-1}(x)$:

$$\theta_m = \arg\min \sum_{i=1}^{N} L(y_i, f_{m-1}(x_i) + T(x_i; \theta_m)) \quad (14)$$

Square error function:

$$L(y, f_{m-1}(x) + T(x; \theta_m)) = [\gamma - T(x; \theta_m)]^2 \quad (15)$$

The residual γ is:

$$\gamma = \gamma - f_{m-1}(x) \quad (16)$$

The boosting tree algorithm is implemented by the forward stepwise algorithm and the additive model, but the optimization of the loss function is more complicated. Friedman developed the gradient boosting algorithm [18]. Its innovation is to approximate the value of the negative gradient of the cost function in the iterative process to the parameters of the regression tree [19], and its negative gradient value is

$$-\left[\frac{\partial L(y, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{m-1}(x)} \quad (17)$$

III. ONLINE MONITORING SYSTEM DESIGN

A. Overall Design

In order to measure the temperature in the aerospace manufacturing process and overcome the shortcomings of the above-mentioned various temperature measurement methods, this paper designs a temperature measurement system. As shown in Figure 1, the monitoring system mainly includes two main parts: the hardware platform layer and the terminal monitoring layer. The hardware platform layer is mainly used to obtain the thermoelectric potential corresponding to the temperature of the working end of the thermocouple and the temperature of the reference end, and realize the data transmission function after processing, that is, the collected data is denoised and passed through a single frequency wireless Wi-Fi module, Transmitted to the terminal host computer monitoring center; the terminal monitoring layer mainly uses the algorithm model to compensate the temperature, and displays the compensated accurate online temperature to achieve high-precision online monitoring of the aerospace manufacturing process. As a result, online temperature measurement of different aerospace components can be realized.
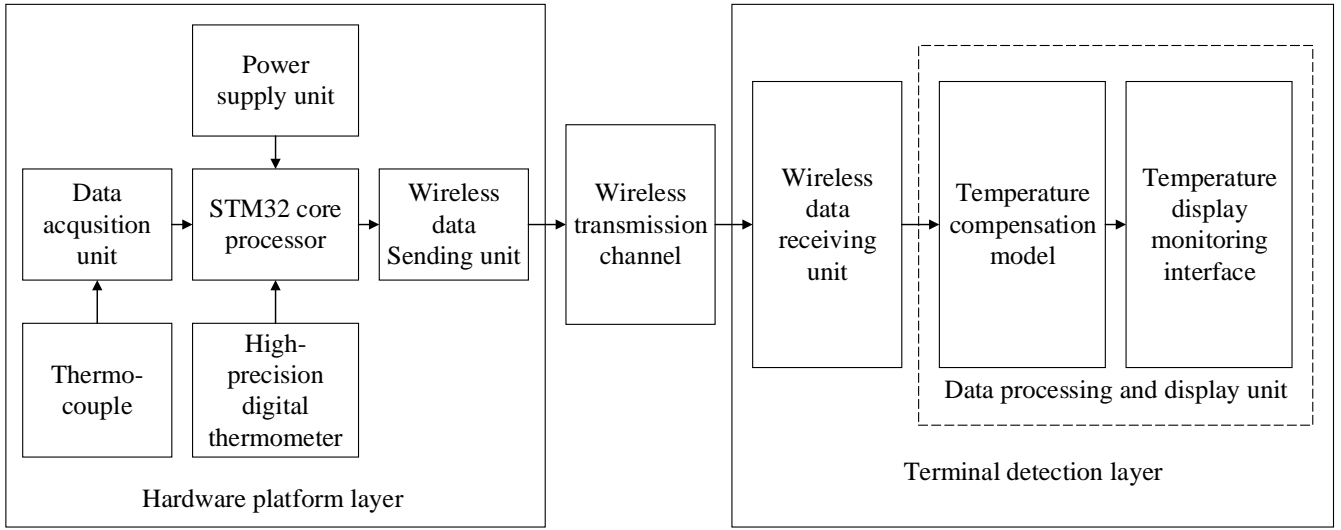
Fig. 1. The overall structure of the monitoring system.

B. Hardware Platform Design

The hardware platform layer is mainly composed of K-type magnetic thermocouple, amplifying and filtering circuit, A/D data sampling circuit, wireless transmission module, high-precision digital temperature sensor, power management circuit and microprocessor. The K-type magneto-thermocouple is connected to the measured object by means of magnetic adhesion to obtain the temperature of the measured object; the amplifying and filtering circuit can amplify the tiny thermoelectric potential signal of the thermocouple, and can effectively filter out the noise under working conditions to avoid noise interference to the compensation algorithm; A/D data sampling circuit converts the analog thermoelectric potential signal processed by amplifier circuit into a discrete digital signal, which is sent and processed by the wireless module; a high-precision digital temperature processor is used to determine the position of the thermocouple's cold end temperature, that is, the ambient temperature is measured, and its high-precision characteristics effectively ensure the accuracy of temperature monitoring; the power management system meets the requirements of different operating voltages such as micro-processing, A/D data sampling circuit and wireless module, and guarantees the operating current's stability and reliability. As shown in Figure 2, the hardware platform structures diagram.
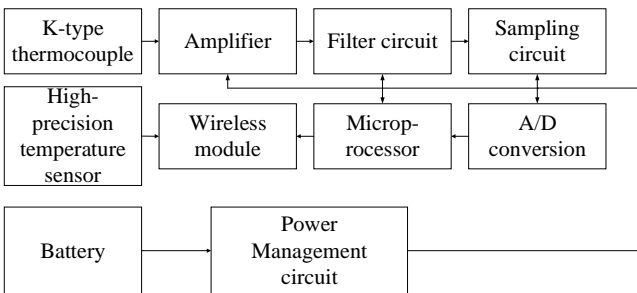

Fig. 2. Hardware platform structure.

C. Hardware Platform Signal Acquisition Software Design

The hardware platform software mainly implements the functions of collecting, processing and sending data of thermoelectric potential data and ambient temperature data. The Wi-Fi wireless transmission chip SX1278 and the sampling chip AD7606 communicate with the microprocessor STM32 via SPI. Figure 3 shows the host computer software interface of the terminal monitoring platform. The software interface is divided into four parts according to functions, which are the wireless serial port basic configuration area, the serial port receiving data display area, the corrected temperature display control area, and the serial port status display area.

Fig. 3. PC software interface of terminal monitoring platform.

Among them, the wireless serial port basic configuration area selects and configures the baud rate, number of digits, stop bit and check digit in the form of a drop-down list. The "open" button triggers the serial port open callback function, and connects after verifying the serial port status. After the same connection is successful, the button displays "close", click again to trigger the serial port to close the callback function; the serial port receiving data display area configures the data display format, selects whether it is hexadecimal, and displays the original data received by the serial port; the temperature after calibration The display control area draws and displays the temperature data after online temperature correction.

IV. DESIGN OF TEST DATA ACQUISITION PLATFORM

In order to realize the temperature monitoring in the aerospace manufacturing process, this paper carries out the following simulation experiment design. In order to realize online temperature monitoring, a temperature control box is used to control the temperature of the cold end, and a high-precision high-temperature thermostat is used to accurately

control the temperature of the heat source, that is, the temperature to be measured. Figure 4 is a prototype diagram of the data acquisition experiment platform, which is composed of a temperature control box, a constant temperature bath, a standard-verified high-precision thermometer, a K-type thermocouple, and amplifying, filtering, and sampling hardware circuits; and the hardware platform transmits the data of cold end temperature and hot end thermoelectric potential to the terminal monitoring upper computer layer.
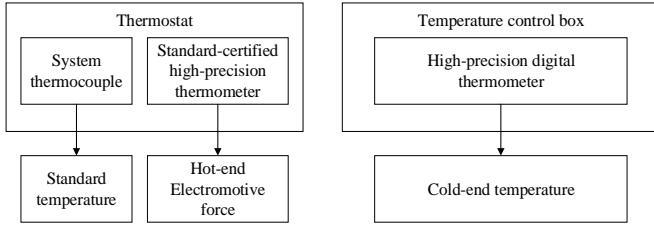
Fig. 4. Experimental platform prototype.

The main device in the test data acquisition platform is a high and low temperature control box. The GPS-3 model was selected. Its controllable temperature range is -70°C-180°C, which can entirely cover the application scenario of this subject, that is, the temperature of the cold end of the thermocouple in temperature monitoring range. For instance, in the key components of the "Gofen No. 4" camera, the temperature is 20°C ± 0.5°C, which requires high temperature stability. Its temperature range is within the scope of this article [20]. "Fluke 1552A-Ex Standard Thermometer" is a high-precision thermometer with standard verification. The temperature measurement range is -80°C-300°C, which can fully meet the temperature monitoring range; the accuracy is ±0.05°C; the response time is 21s, and the measured value is used as the standard temperature; The ST-GHT-V5 high-temperature thermostat of "Liyida" has a monitoring range of ambient temperature +30°C—300°C. It uses PID high-precision program control to achieve precise temperature control within the range of ±0.01°C.

The sample data of the temperature compensation algorithm model is predominantly obtained by the test data acquisition platform built in the laboratory, as shown in Figure 30. The constant temperature bath is a high-precision and high temperature constant temperature bath with a controllable temperature of 60 °C-300 °C, which is used to simulate the temperature of the working end obtained when the thermocouple of the monitoring system adheres to the object to be measured. The temperature control box is a temperature box with a controllable temperature of -70°C to 180°C to simulate the ambient temperature of the cold end of the monitoring system. The high-precision digital temperature sensor DS18B20 obtains the ambient temperature of the temperature control box with an accuracy of ±0.5°C, which is used as the temperature value of the cold junction of the thermocouple; in addition, the high-precision thermometer that has passed the standard verification and the system thermocouple are in the "heat source high-precision high-temperature constant temperature bath". The output temperature of the high-precision thermometer is used as the standard value of the temperature measured in the experiment.

The specific experiment process is controlling the temperature control box in the range of -20°C-40°C, with 2°C or 3°C as the interval, and the high-temperature thermostat in the range of 70°C-275°C, with the interval of 5°C, carry out the temperature test. At the same time, record the temperature output data of the thermometer corresponding to the standard-verified high-precision thermocouple.

The standard temperature, hot-end thermoelectric potential and cold-end temperature data together constitute the data set of the compensation algorithm. Use the test data collection platform experimental data set to train and verify the algorithm model. First, determine that the input X of the model is the cold end temperature T and the thermoelectric potential V; the output Y is the temperature Y of the object under test. It mainly uses regression algorithms for model training, and plans to use regression algorithms such as SVM and GBDT to compare and optimize the algorithm model.

Table 1 shows part of the experimental data obtained by the data acquisition platform, including the measured temperature of the thermostat, the measured thermoelectric potential after amplification and filtering, and the measured thermostatic bath temperature.

TABLE I.   EXPERIMENTAL DATA TABLE (PARTIAL)

Actual temperature of thermostat/°C	Measured thermoelectric potential/ mV	Measured temperature of constant temperature bath/°C	Actual temperature of thermostat/°C	Measured thermoelectric potential/ mV	Measured temperature of constant temperature bath/°C
-18.0	808	70	-8.2	720	75
-18.0	855	75	-8.2	768	80
-18.0	903	85	-8.2	817	85
-18.0	951	90	-8.2	865	90
-18.0	998	95	-8.2	913	95
-13.2	770	75	-3.5	687	75
-13.2	817	80	-3.5	736	80
-13.2	865	85	-3.5	784	85
-13.2	913	90	-3.5	832	90
-13.2	962	95	-3.5	881	95

The measured temperature of the thermostat and the thermoelectric potential are used as independent variables, and the measured constant temperature bath temperature is used as the dependent variable to match the model. Thermocouple temperature test data has a large variance, while the GBDT algorithm model performance has a small relationship with the value of the training sample. Only need to consider the division boundary, such as ID3 algorithm, C4.5 algorithm, and CART algorithm. The splitting of the tree is only related to information entropy and Gini. The coefficient is related. Therefore, the gradient boosting model is considered to directly train the original experimental data.

Based on the measured temperature of the thermostat, the thermoelectric potential of the measured temperature, and the measured temperature of the constant temperature bath, the GBDT model of static temperature correction is trained; the measured temperature is corrected based on this model. The process is shown in Figure 5:

V. EXPERIMENT AND RESULT ANALYSIS

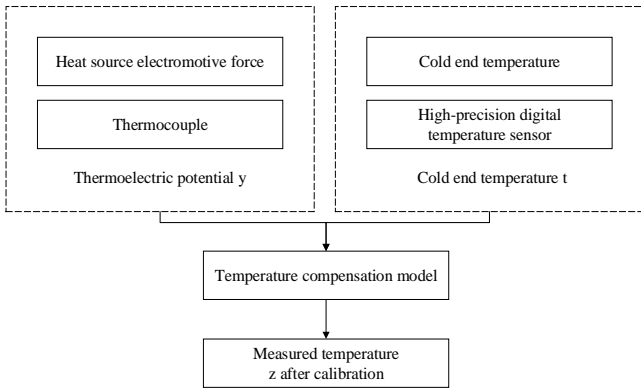### A. Comparison of Static Error Experiment of Online Monitoring System

Fig. 5. System regression prediction flowchart.

The following experiments simulate different methods of temperature measurement in the aerospace manufacturing process, and verify the performance of the GBDT model. Static error comparison of LR, SVR and GBDT algorithms under experimental conditions. Based on the Python language and Scikit-learn package on the Windows platform, the LR, SVM, and GBDT models are trained using the data of the thermocouple test. The data set obtained by the data acquisition platform, the training data sample is shown in Table 2, a total of 1968 items. The record includes the actual measured temperature of the thermostat—the cold end ambient temperature and the thermoelectric potential. The prediction target is the measured temperature of the constant temperature bath.

TABLE II.   TRAINING SET DATA TABLE (PARTIAL)

Actual temperature of thermostat/°C	Measured thermoelectric potential/mV	Measured temperature of constant temperature bath/°C	Actual temperature of thermostat/°C	Measured thermoelectric potential/mV	Measured temperature of constant temperature bath/°C
-18	808	70	-8.2	720	75
-18	855	75	-8.2	768	80
-18	903	85	-8.2	817	85
-18	951	90	-8.2	865	90
-18	998	95	-8.2	913	95

The parameters of the support vector machine (SVR) are set as follows: the penalty factor is 0.1, the kernel function is a linear function, the gamma coefficient is set to 0.001, and the MAE and MSE on the training set are 1.4563 and 8.6782. The parameters of the GBDT algorithm are set to 600 basic trees, the maximum depth is 7, the minimum sample for decision tree division is set to 2, and the learning rate is 0.1. The MAE and MSE on the training set are 0.4716 and 0.3528. The results are shown in Table 3.

TABLE III.   MODEL EVALUATION INDEX TABLE

Model	EV	MAE	MSE	R^2
LR	0.9935	1.4036	8.3852	0.9972
SVR	0.9971	1.4563	8.6782	0.9971
GBDT	0.9999	0.4716	0.3528	0.9999

Table 4 shows the comparison of the measured values of the measured constant temperature bath temperature after calibration of different models under the laboratory test data collection platform.

TABLE IV.   COMPARISON TABLE OF MEASURED VALUES AFTER CALIBRATION OF DIFFERENT MODELS

Experiment number	Constant temperature bath temperature/°C	LR correction/°C	SVR correction/°C	GBDT correction/°C
1	70	68.62	68.52	69.52
2	82	80.63	80.74	81.57
3	94	92.89	95.93	94.26
4	106	104.71	107.62	105.55
5	118	116.62	116.64	118.37
6	130	128.47	128.71	129.51
7	142	140.81	143.36	141.72
8	154	152.79	152.89	153.58
9	166	164.01	167.54	166.29
10	178	176.72	176.38	178.47
11	190	188.35	191.55	189.46
12	202	203.71	201.05	202.51
13	214	215.06	215.07	213.46
14	226	227.70	224.62	226.57
15	238	239.34	239.46	237.85
16	250	251.39	248.57	250.17
17	262	263.45	263.87	261.36
Mean absolute error		1.42	1.43	0.45

Integrating the model average absolute error MAE and parameter indicators such as Figure 7, it can be concluded that the predicted value of the GBDT model fits best with the true value, and the error on the experimental data set is the smallest. It can be seen from Table 4 that compared to the LR model and the SVR model, the GBDT model is easier to fit the test data in the thermocouple static temperature measurement, and the temperature correction effect is better.

B. Dynamic Error Experiment Comparison of Online Monitoring System

Under experimental conditions, the simulated measured temperature is input in the form of a step, and the comparison of the dynamic errors of the monitoring system is observed. Taking the data acquisition platform as the experimental object, choose the optimal algorithm for compensating static errors—optimized and tuned Xgboost as the thermocouple correction algorithm, and compare the difference in measured values between adding a dynamic compensator and not adding a dynamic compensator. Specifically, taking 80°C to 90°C as an example, the temperature control box that controls the temperature of the cold end in the data acquisition platform is set to 27°C, two high-precision constant temperature baths are selected, and the temperature of constant temperature bath 1 is set to 80°C. The temperature of the thermostatic bath 2 is set to 90°C and stabilized for more than 30 minutes. Fix the temperature of the cold end temperature control box, place the thermocouple in thermostatic bath 1 for 20 minutes, then quickly place it in thermostatic bath 2, and record the temperature data of the whole process. Table 5 is a comparison table of specific temperature measurement values every 2s.

Time/s	Xgboost correction temperature/°C	Xgboost correction temperature absolute error/°C	Xgboost+ PSO correction temperature /°C	Xgboost+ PSO correction absolute temperature error/°C
2	80.2	9.8	82.9	7.1
4	80.6	9.4	86.5	3.5
6	81.4	8.6	89.0	1.0
8	82.2	7.8	89.9	0.1
10	83.0	7.0	90.1	-0.1
12	83.9	6.1	90.1	-0.1
14	84.7	5.3	90.0	0.0
16	85.5	4.5	90.0	0.0
18	86.3	3.7	90.0	0.0
20	86.9	3.1	90.0	0.0
22	87.5	2.5	90.0	0.0
24	88.0	2.0	90.0	0.0
26	88.5	1.5	90.0	0.0
28	88.8	1.2	90.0	0.0
30	89.2	0.8	90.0	0.0
32	89.4	0.6	90.0	0.0
34	89.6	0.4	90.0	0.0
36	89.7	0.3	90.0	0.0
38	89.9	0.1	90.0	0.0
40	90.0	0.0	90.0	0.0

It can be seen from Figure 8 and Table 5 that under the excitation of the step signal, the output of the magnetic thermocouple system and the output of the compensated system, the magnetic thermocouple monitoring system purely statically corrected by Xgboost has a relatively noticeable time lag, and the response time of the system. It is about 40s, and the response time of the system after adding the dynamic compensator is significantly improved, shrinking to about 14s, which can reduce the dynamic errors.

VI. CONCLUSION

This paper discusses the temperature measurement part of the aerospace manufacturing process and implements the method of thermocouple as the temperature measuring element, and establishes the correction of the temperature measurement process in aerospace manufacturing based on the gradient boost algorithm in machine learning, and finally completes the development of the temperature online monitoring system based on the gradient boost algorithm. The core is to use algorithms to compensate for the effects of various error factors of thermocouples. The overall structure of the system includes two parts: a hardware platform and a terminal monitoring platform. Among them, the hardware platform collects the thermoelectric potential data and cold end temperature data of the thermocouple, and communicates with the terminal monitoring platform through wireless Wi-Fi to realize data transmission; the terminal monitoring layer uses the gradient boost algorithm to perform nonlinear correction of the thermocouple. As well as cold end compensation, a dynamic compensator is used to compensate for dynamic errors, and the temperature after correction and compensation is displayed online to achieve accurate monitoring. It provides a certain reference value in the field of temperature detection in aerospace manufacturing in the future.

REFERENCES

[1] Liu Gang, Li Zhengju, Zhu Xi, Xie Weihua and Li XiangYang," Closed-loop temperature control of high-power telecommunicationsatellite in thermal vacuum test, "Spacecraft Environment Engineering, 2019 Issue 4 369-373. [In Chinese]

[2] HU Xiao-feng, CHEN Hong-yu, FAN Ya-jie, WANG Qiang, JI Qi-zheng, PENG Hao, XIE Xi-ning and ZHANG Quan, "Influences of Temperature on Static Discharge Pattern of GEO Orbit Solar Array"Equipment Environmental Engineering, 2020 Issue 06, Pages: 86-94. [In Chinese]

[3] Yin Yazhou, Fu Weichun, Lu QingRong and Qian ZhiYing, "Thermal Design of GF-7 Satellite Based on Dimensional Stability, " Spacecraft Engineering, 2020-03 Issue, Pages: 82-88. [In Chinese]

[4] Tang Zhaoting, "The Design of Spacecraft Thermocouple Thermometer, " Xi'an : Xidian University, 2011. [In Chinese]

[5] Biao Z and Jun S,"Research and Design of Cold Junction Compensation in Thermocouple Temperature Measurement," Journal of Shanghai Ship & Shipping Research Institute, 2013, 36(1): 50-53. [In Chinese]

[6] Hai-Tao Z, Shan L and Tao G, "Improvement of Cold-conjunction Compensation of Thermocouple,"Instrument Technique and Sensor, 2011(7): 11-14. [In Chinese]

[7] Shangchun Fan, "Sensor technology and application," Beihang University Press, 2010(7):208-216. [In Chinese]

[8] Qu Jiaqi, Qian Zheng, Li Yutao and Tian Hongwei, "Research on online temperature monitoring method of thermocouple based on GBDT," Electrical Measurement & Instrumentation, in press. [In Chinese]

[9] Zeng XiaoXin and Qiu LiYun, "Influence Factors and Optimization Method of Thermocouple Measurement Accuracy,"Automation & Instrumentation , 2016 Issue 09, Pages: 37-40. [In Chinese]

[10] Joseph D, Collins S, "Modeling, calibration, and correction of nonlinear illumination-dependent fixed pattern noise in logarithmic CMOS image sensors," IEEE Transactions on Instrumentation and Measurement, 2002, 51(5): 996-1001

[11] Zhou ZhiHua, "Machine Learning,"Tsinghua University Press, 2016(1):106-107. [In Chinese]

[12] Agrawal R. J. and Shanahan J. G., "Location disambiguation in local searches using gradient boosted decision trees," Sigspatial International Conference on Advances in Geographic Information Systems[C]. ACM, 2010: 129-136.

[13] Safavian S. R. and Landgrebe D., "A survey of decision tree classifier methodology," IEEE Transactions on Systems, Man and Cybernetics, 1991, 21(3):660-674.

[14] Yang XueBing, Zhang Jun, "Decision Tree and Its Key Techniques,"Computer Technology and Development, Issue 01, 2007 Pages: 43-45. [In Chinese]

[15] Zhu XiaoLiang, Yan HongCan., Wang Jian and Wu ShangZhuo, "Research and application of the improved algorithm C4.5 on Decision tree," 2009 International Conference on Test and Measurement[C], Hong Kong, 2009: 184-187.

[16] Burrows, William R. Benjamin, Mario Beauchamp, Stephen Lord, Edward R. McCollor, Douglas Thomson, Bruce, "CART Decision-Tree Statistical Analysis and Prediction of Summer Season Maximum Surface Ozone for the Vancouver, Montreal, and Atlantic Regions of Canada". Journal of Applied Meteorology, 2010, 34(8):1848-1862.

[17] K A Grajski, L Breiman, G Viana Di Prisco and W J Freeman,"Classification of EEG Spatial Patterns with a Tree-Structured Methodology: CART,"IEEE Transactions on Biomedical Engineering, 2007, 33(12): 1076-1086.

[18] Friedman J. H., "Greedy Function Approximation: A Gradient Boosting Machine," The Annals of Statistics, 2001, 29(5):1189-1232.

[19] Yang Can, "Road network extraction from remote sensing images based on XGBoost"Microcomputers and Applications ,2017, Issue 24, pages 28-31. [In Chinese]

[20] Yu Feng, Xu Nana, Zhao Zhenming, "Orbit temperature data analysis and thermal design optimization for space camera on GF-4 satellite" Journal of Beijing University of Aeronautics and Astronautics, ISSN 1001-5965, CN 11-2625/V. [In Chinese]

A Temperature and Humidity Compensation Method for On-board NOx Sensors with LSTM Network

Anran Huang
School of Automation Science and Electrical Engineering
Beihang University
Beijing, China
huanganran@buaa.edu.cn

Yingming Lyu
Research Institute of Technology Development
Beijing Machinery Industry Bureau
Beijing, China
lym928@163.com

Zicheng Guo
School of Automation Science and Electrical Engineering t
Beihang University
Beijing, China
zgzc1221@163.com

Xiangyang Zhao*
School of Automation Science and Electrical Engineering
Beihang University
Beijing, China
zhaoxiangyang@buaa.edu.cn

Abstract—**NOx is a pollutant component of vehicle exhaust. On-board NOx sensors usually operate under electro-chemical reactions. Temperature and humidity can influence these reactions, resulting in the inaccuracy of measurements. Thus, a method based on long short-term memory (LSTM) network for temperature and humidity compensation of on-board NOx sensors is proposed in this paper. The LSTM network is trained with the sensor-measured and the true NOx concentrations alongside with the temperature and relative humidity of tested gas. The testing results of the LSTM network shows that this method has good performance for the temperature and humidity compensation, and can efficiently improve the accuracy of on-board NOx sensors.**

Keywords—LSTM network, NOx sensors, compensation, on-board measurement

I. INTRODUCTION

NOx ($NO + NO_2$) is toxic to human body and atmosphere. The inhalation of NOx from respiratory tract will damage the lung structure and cause acute poisoning. When emitted into the atmosphere, NOx can lead to the formation of acid rain and will form photochemical smog when reacted with hydrocarbons. It can also do damage to the tropospheric ozone (O_3), resulting in the aftermath of global warming [1].

NOx can be produced in the combustion of fossil fuel under high temperature. This procedure often happens in engines of vehicles and aircrafts, internal-combustion engines and industrial furnaces. Among all emission sources, the vehicles powered by gasoline and diesel are the greatest contributors. The vehicle source in China can take up more than 30% of anthropogenic NOx emissions, making it the second largest source after industrial emissions. According to the China National Development and Reform Commission, the current vehicle ownership of China is approximately estimated as 350 million and is still rapidly growing. Although it indicates the prosperity of the nation, this astonishing number also represents the risk of a huge NOx emission amount. Surprisingly also, diesel vehicles account for 71.2% of total vehicle NOx emissions. Meanwhile, only 9.1% of the vehicles in China are diesel-powered [2]. With the unignorable pollutions from NOx and the growing trend of NOx emissions from vehicles, much importance has been attached to the reduction of vehicle NOx emissions worldwide.

Early in 1992, the EU published the vehicle emission standard Euro 1 where NOx and HC emission amounts were calculated together and regulated. In 2000, NOx was listed as a separate pollutant for the first time in standard Euro 3, and its maximum emission was regulated as 0.5g/km for diesel vehicles and 0.15g/km for gasoline ones. The Euro 6 standard in operation now rectifies an upper limit of NOx emission as 0.08g/km for diesel vehicles and 0.06g/km for gasoline [3]. It is already announced that the EU will operate the Euro 7 standard, with an even stricter limit of NOx emission amount, in 2026. In China, the emission standard took a relatively backward start but is catching up quickly with the standard Euro 6. The first vehicle emission standard of China was carried out in 2001. The standard was equivalent with Euro 1. Since then, the Chinese Central Government has been updating the standard quickly with efficient policies to reduce the NOx emission amount from vehicles. At present, the first stage of the current emission standard China VII has just been implemented. In this standard, the maximum NOx emission is 0.06g/km for both diesel and gasoline vehicles [4]. The stricter standards on vehicle NOx emission requires better accuracy in on-board NOx measurements. Thus, methods to rectify the data from NOx sensors should be proposed.

The principle of NOx sensors is usually based on electro-chemical reactions. The most common reaction in NOx sensors is redox reaction. Via the reactions, the concentration of NOx will be converted into electrical signals. The characteristics of electrochemical reactions enable NOx sensors to operate with precision, convenience and immediacy, but it will also result in the dramatical drift on accuracy of the sensors with temperature and humidity [5]. Although many NOx sensors are equipped with interior self-compensators, it is still recommended that they work under standard environment. But the working environment for on-board NOx sensors is often too harsh to rely on self-compensators only. Hence, an external compensation method for the sensors is still necessary.

The traditional method to compensate the drifts for sensors is hardware compensation. This method requires exterior circuits that are designed and connected to sensors. These circuits have limited precision and are complicated to design. Besides, they require exquisite calibration process. For these drawbacks, compensation via software has been given rising attention. With optimization algorithms and neural networks,

---

*Xiangyang Zhao is the corresponding author. (e-mail: zhaoxiangyang@buaa.edu.cn).
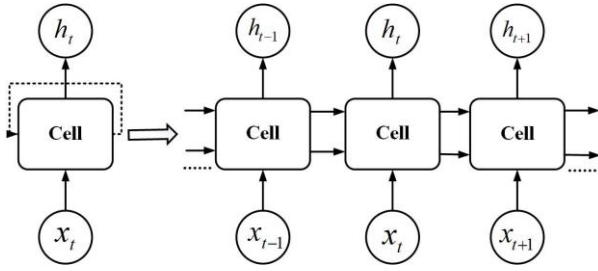
Fig. 1. The recurrent structure of LSTM

compensation methods are becoming more universal and accurate. Wu et al. [6] use Hammerstein neural network on infrared thermometer sensor for dynamic error compensation. Patra et al. [7] propose a method to apply neural networks to compensate the nonlinear environmental influence on sensors and test the performance on a capacitive pressure sensor. De Meng et al. [8] build a model of hysteresis based on neural network, and apply the model to compensate the piezoelectric stack actuator on helicopter. The results of these studies show the capability of artificial neural networks to fulfill the tasks of sensor compensation. Based on these achievements, the application of neural networks for the temperature and humidity compensation for on-board NOx sensors seems practical and promising.

The research work in this paper presents a method with long short-term memory (LSTM) network to compensate the errors from temperature and humidity on NOx sensors. The goal is to reduce the error between output NOx concentration and the real value to less than 2% according to engineering requirements. An on-board NOx sensor and temperature and humidity sensors were utilized in the experiments. The tested gas was generated with certain concentration. The data obtained were used for the training and testing of the network.

II. ESTABLISHMEN OF LSTM COMPENSATION MODEL

A. LSTM Network Model

LSTM network [9] is a neural network model improved on the basis of recurrent neural network (RNN). It has the typical recurrent structure of RNN and can be expanded into chain structure too, as is shown in Fig. 1. At time t, the input of the network is x_t, and h_t serves as the output state of the cell. Like RNN, h_t also is the hidden state of the cell.

It is notable that, from Fig.1, each cell outputs 2 states to the next cell. In fact, the main characteristic that differs LSTM network from conventional RNN is this structure of its cells. Fig. 2 shows the inner structure of an LSTM cell. Apart from h_t, another hidden state c_t that passes to the next cell is calculated. This hidden state is called "cell state". Instead of merely activation functions, sigmoid 1, 2, 3 and tanh 1 in Fig.2 represent 4 independent feedforward network layers. The network layers, of course, adopt their corresponding activation functions. The numbers of the hidden units in these networks are uniform and are determined by a parameter called "hidden size" of LSTM network.

There are 3 stages of calculation in an LSTM cell. The first "forget" stage will selectively neglect the information from the previous cell and outputs a vector g_t. The second stage "selective memory" will partially memorize the input information. Apart from its output vector r_t, this stage will produce another output ς_t as a candidate cell information to update c_t. The final "output" stage will calculate the value for h_t and pass it to the outside and to the next cell. At this stage, a vector p_t will be calculated. Also, the update of the value for c_t will happen between the second and third stage. This is because the value of c_t will consequently impact h_t.

For each stage, a distinct "gate" structure is involved. A "gate" is composed of a sigmoid feedforward network and a pointwise product operation, as is shown in the dotted boxes in Fig.2 . Since a sigmoid layer produces an output between 0 to 1, it can actually decide the amount of information that passes through the gate. When it "decides" that no information shall pass, its output will be 0. When all information can pass through the gate, its output is 1. According to the sequence of the stages, the relevant gates are named as "forget gate", "input gate" and "output gate". In fact, the "gates" qualify LSTM to have long-term memory.

With this unique cell structure, LSTM network can avoid the long-term dependencies which often happen when training RNN. This feature protects LSTM network from having vanishing gradients and allows its superior performance over RNN.

B. Analysis of LSTM Network for Compensation

Based on the task of LSTM network in this research, the input of the network can be defined as a sequence of three-dimensional vectors. The elements of each vector are the
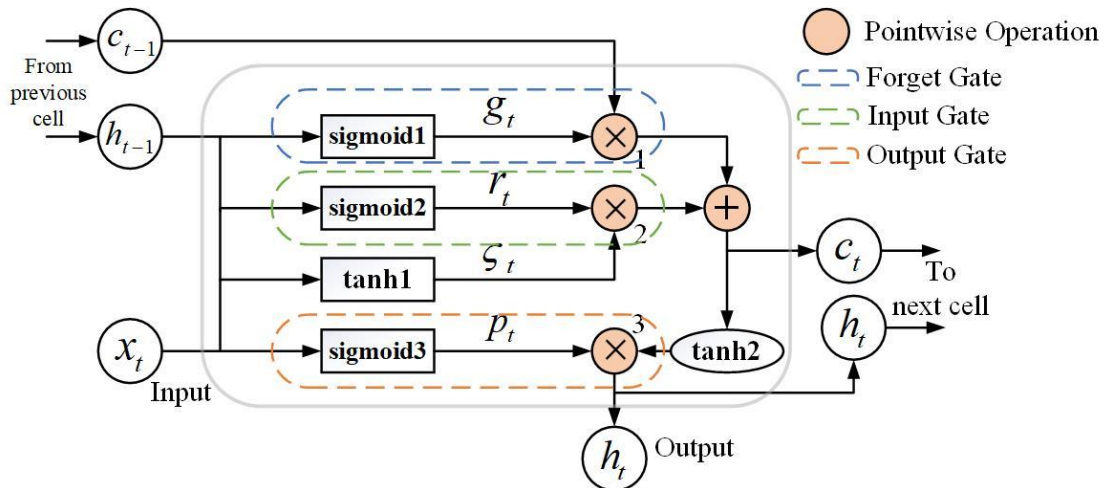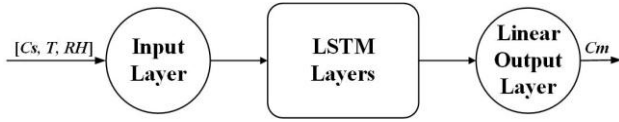

Fig. 2. The cell structure of LSTM

Fig. 3. The complete network structure

output concentration of NOx sensor, the temperature of the tested gas and the relative humidity of the gas. They are respectively denoted as *Cs*, *T* and *RH*. Correspondingly to the vector, the output of the network is the modified NOx concentration denoted as *Cm*. The true concentration C_t matching with the input vector serves as the expected value of training .

Since the temperature and humidity compensation of on-board NOx sensors essentially is a regression problem, the output layer should be a linear layer. Accordingly, the network is structured as Fig. 3. In this network, there are 2 LSTM layers. The hidden size of the network is 20.

The dimensions of hidden states h_t and c_t are in correspond with hidden size. For each feedforward network in the cell shown in Fig.2, their weights and biases are denoted as [W_{s1}, b_{s1}], [W_{s2}, b_{s2}], [W_t, b_t] and [W_{s3}, b_{s3}] from top to bottom. With all parameters defined, the updates of hidden states in the cell can be derived.

When a vector $x_t=[C_{st}, T_t, RH_t]$ is input to the cell, it will firstly pass through the forget gate. At forget gate, the computational process is as below:

$$g_t=sigmoid(W_{s1}*[h_{t-1},x_t]+b_{s1}) \qquad (1)$$

In (1), the output g_t and h_{t-1} are 20-dimensional vectors. Therefore, W_{s1} should be a 23*20 matrix and b_{s1} a 20-dimensional row vector.

Then at the input gate, the process to selectively memorize new information goes as (2), and the generating process of ς_t is shown in (3):

$$r_t=sigmoid(W_{s2}*[h_{t-1},x_t]+b_{s2}) \qquad (2)$$

$$\varsigma_t=tanh(W_t*[h_{t-1},x_t]+b_t) \qquad (3)$$

Likewise, both r_t and ς_t have 20 dimensions. Thus, the dimensions of W_{s2}, W_t are consistent with that of W_{s1}. So are the dimensions of b_{s2} and b_t with b_{s1}.

After obtaining g_t, r_t and ς_t, the cell state c_t can be now updated according to the following equation:

$$c_t=g_t \odot c_{t-1} + r_t \odot \varsigma_t \qquad (4)$$

where $\odot$ represents Hadamard product between matrices.

Equation (4) shows how the current cell state is composed of : g_t from forget gate filters the cell state of the previous cell, and r_t from input gate partly retains the candidate cell state produced in the current cell.

The last step is to calculate the value of output ht. At this stage, the output gate produces a vector p_t:

$$p_t=sigmoid(W_{s3}*[h_{t-1},x_t]+b_{s3}) \qquad (5)$$

Equation (5) can be interpreted as a condition to analyze the cell output h_t based on h_{t-1} and x_t. In (5), the dimensions of the vectors are the same as those correspondingly from (1) and (2).

With the previous work, h_t can be finally computed as:

$$h_t=p_t \odot tanh(c_t) \qquad (6)$$

The equations above automatically complete the forward propagation in an LSTM cell. Afterwards, a back propagation process is carried out to update all the weights. Finally, the network will give a prediction of the true NOx concentration from the sensor's measurement, the temperature and relative humidity.

III. MODEL TRAINING AND TESTING

Before training the network model, the concentration data of NOx under different temperature and humidity conditions should be collected. The concentrations are measured in ppm (parts per million). Under the standard environment (20 °C, 60%RH), clean NOx is diluted into required concentrations. The temperature and relative humidity of the gas are controlled to various set values.

The sensor adopted in the experiment is NX1 CiTiceL by City Technology. It is an on-board NOx sensor developed for vehicles. The NOx sensor has an internal filter for SO_2 to avoid the cross interference. Via a data acquisition board, the outputs of the sensor are transferred from analogue to digital.

The LSTM network is established, trained and tested on Pytorch, a popular framework for deep learning models.
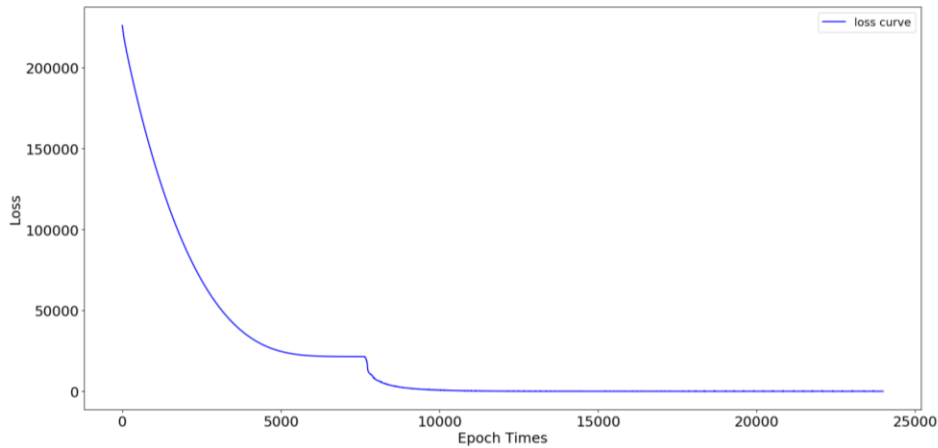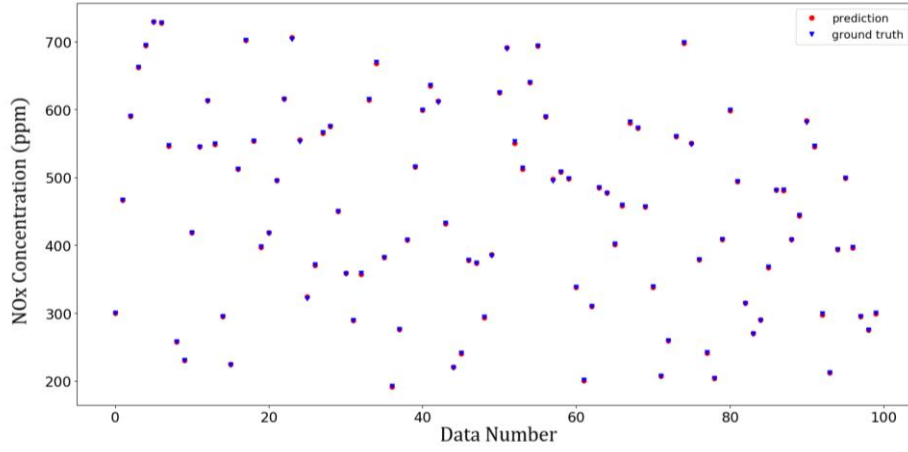

Fig. 4 Loss curve of LSTM network

Fig. 5 The prediction and ground truth of testing set

A. Training of LSTM Network

Without pre-training of weights, the parameters of the LSTM network model are trained until the loss coverages. We define the losses of the LSTM network as the errors between the true NOx concentrations and predicted concentrations. The losses are calculated in the form of mean square error (MSE) as (7):

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(c_p^{(i)} - c_r^{(i)}\right)^2 \tag{7}$$

where N represents the total number of data, c_p denotes the predicted NOx concentrations and c_r denotes the true NOx concentrations. MSE automatically converts the losses to their absolute values and emphasizes the impacts of data with greater losses during training.

The optimization algorithm for the network is Adam. It is an adaptive algorithm to update the weights. Adam shows superiority over other optimization methods and has wide application in deep learning field.

The dataset contains 1,000 pairs of collected data. It is grouped randomly into a training set with 900 pairs of data and a testing set with 100 pairs.

B. Results of Compensation

The performance of the LSTM network for compensation can be demonstrated from several perspectives such as the loss of the network, the accuracy of compensation and the average relative error of the network. Fig. 4 shows the loss curve when training the network. Fig. 5 shows the predicted values of NOx concentrations and true values of the testing set. The relative error curve between them for each set is plotted in Fig. 6. The average relative error of the testing set is 0.149%. These figures provide visible evidences for the performance of the network.

Due to the close gaps between the predicted values and true values in Fig. 5, another 2 graphs are drawn for a legible view of the results. The elements in the testing set are filtered by the range of true NOx concentrations in to a data group.
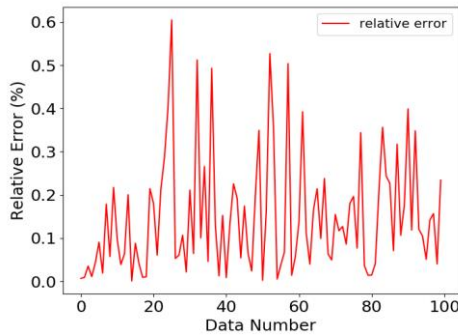
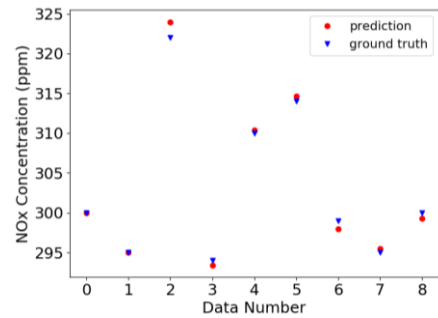Fig. 6 The relative errors of testing set

Fig. 7 The prediction and ground truth concentration of data group I

TABLE I.   INFORMATION OF DATA GROUP I

Parameter	Data Number								
	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Temperature (℃)	30	27	50	42	41	31	36	20	40
Relative Humidity (%)	50	42	59	42	61	60	50	60	50
Measured Concentration (ppm)	307	297	344	307	330	326	311	297	317
True Concentration (ppm)	300	295	322	294	310	314	299	295	300
Predicted Concentration (ppm)	299.98	294.99	323.95	293.42	310.34	314.66	297.96	295.46	299.30
Relative Error	0.007%	0.003%	0.605%	0.197%	0.110%	0.210%	0.348%	0.156%	0.233%

596

Parameter	Data Number							
	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Temperature (℃)	40	39	33	37	30	20	40	50
Relative Humidity (%)	40	48	42	39	60	61	55	40
Measured Concentration (ppm)	604	633	626	628	611	613	650	627
True Concentration (ppm)	590	613	615	615	600	611	625	599
Predicted Concentration (ppm)	590.21	613.39	616.30	614.38	600.05	612.38	625.02	598.91
Relative Error	0.036%	0.064%	0.211%	0.101%	0.008%	0.226%	0.003%	0.015%

Then, we plot the predicted and true NOx concentration values of the group. In Fig. 7, the true concentration range of data group I is from 290 to 330 ppm. Table I displays the relative parameters of the data group in Fig. 7.

The 9 testing data sets in data group I are numbered from 0 to 8. Their predicted and true values of NOx concentration is plotted. The largest relative error in this group is 0.605%.

In Fig. 8, the range of the true concentrations to filter data group II is changed to 590~630 ppm. Table II presents the parameters for data group II of Fig. 8.

The 8 testing data sets in data group II are numbered from 0 to 7. Fig. 8 demonstrates the prediction and ground truth values of the NOx concentration. In this group, the largest relative error is 0.226%.

The results of the testing process shows that the LSTM network we built is able to compensate the temperature and humidity influence for on-board NOx sensors and can improve the accuracy of measurements.

IV. CONCLUSIONS

In this paper, a temperature and humidity compensation method with LSTM network for on-board NOx sensors is proposed. The characteristics of the compensation model based on LSTM network is analyzed, and the principle of the model is discussed with deduction procedures. We collected the data of temperature, humidity, sensor-measured NOx concentrations and the true concentrations for the network. Through training, the LSTM compensation network shows positive performance of compensation in the process of testing. The results prove that the compensation method for on-board NOx sensors with LSTM network can compensate the drift caused by temperature and humidity, so as to improve the measurement results and monitor the NOx concentrations in vehicle exhaust better.
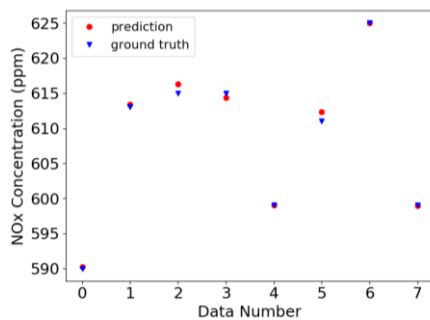

Fig. 8 The prediction and ground truth concentration of data group II

REFERENCES

[1] Duncan, Bryan N., Lok N. Lamsal, Anne M. Thompson, Yasuko Yoshida, Zifeng Lu, David G. Streets, et al. "A Space-Based, High-Resolution View of Notable Changes in Urban NOx Pollution around the World (2005-2014)." Journal of Geophysical Research: Atmospheres 121, no. 2 (2016): 976–96.

[2] Jiang, Lei, Yuan Chen, Haifeng Zhou, and Shixiong He. "NOx Emissions in China: Temporal Variations, Spatial Patterns and Reduction Potentials." Atmospheric Pollution Research, 2020. https://doi.org/10.1016/j.apr.2020.06.003.

[3] Fuć Paweł, Maciej Siedlecki, Natalia Szymlet, Barbara Sokolnicka, Łukasz Rymaniak, and Michał Dobrzyński. "Exhaust Emissions from a EURO 6c Compliant PC Vehicle in Real Operating Conditions." Journal of KONBiN 49, no. 4 (2019): 421–40.

[4] Gu, Xingke, Shasha Yin, Xuan Lu, Huan Zhang, Lingling Wang, Ling Bai,et al. "Recent Development of a Refined Multiple Air Pollutant Emission Inventory of Vehicles in the Central Plains of China." Journal of Environmental Sciences 84 (2019): 80–96.

[5] Lin, Jun, Huan Gao, Xiufeng Wang, Chunyan Yang, Yi Xin, and Xianfeng Zhou. "Effect of Temperature on the Performance of Electrochemical Seismic Sensor and the Compensation Method." Measurement 155 (2020): 107518.

https://doi.org/10.1016/j.measurement.2020.107518.

[6] Wu, Dehui, Songling Huang, Wei Zhao, and Junjun Xin. "Infrared Thermometer Sensor Dynamic Error Compensation Using Hammerstein Neural Network." Sensors and Actuators A: Physical 149, no. 1 (2009): 152–58.

[7] Patra, Jagdish Chandra, Ee Luang Ang, Amitabha Das, and Narendra Shivaji Chaudhari. "Auto-Compensation of Nonlinear Influence of Environmental Parameters on the Sensor Characteristics Using Neural Networks." ISA Transactions 44, no. 2 (2005): 165–76.

[8] Meng, De, Pinqi Xia, Kai Lang, Edward C. Smith, and Christopher D. Rahn. "Neural Network Based Hysteresis Compensation of Piezoelectric Stack Actuator Driven Active Control of Helicopter Vibration." Sensors and Actuators A: Physical 302 (2020): 111809. https://doi.org/10.1016/j.sna.2019.111809.

[9] Horchreiter, Sepp, and Jürgen Schmidhuber. "Long Short-Term Memory." Neural Computation 9, no. 8 (December 1997): 1735–80.

[10] Chen, Fu, and Fan Duo. "The Prediction of ENSO Indexes Based on Time Series LSTM Model." Climate Change Research Letters 08, no. 03 (2019): 287–95.

[11] Li, Yan, and Yifei Lu. "LSTM-BA: DDoS Detection Approach Combining LSTM and Bayes." 2019 Seventh International Conference on Advanced Cloud and Big Data (CBD), 2019.

A Three-Dimensional Metamaterial Resonator in Low Terahertz Frequency

Shengnan Li
Nanshan District, Science and Technology Park, Xi'an Jiaotong University Shenzhen Academy, State Key Laboratory for Manufacturing Systems Engineering Xi'an Jiaotong University
Xi'an, China
lsn3119101110@stu.xjtu.edu.cn

Zhonglei Shen
Nanshan District, Science and Technology Park, Xi'an Jiaotong University Shenzhen Academy, State Key Laboratory for Manufacturing Systems Engineering Xi'an Jiaotong University
Xi'an, China
zlshen@stu.xjtu.edu.cn

Donghai Han
Nanshan District, Science and Technology Park, Xi'an Jiaotong University Shenzhen Academy, State Key Laboratory for Manufacturing Systems Engineering Xi'an Jiaotong University
Xi'an, China
donghai19@stu.xjtu.edu.cn

Liuyang Zhang[*]
Nanshan District, Science and Technology Park, Xi'an Jiaotong University Shenzhen Academy, State Key Laboratory for Manufacturing Systems Engineering Xi'an Jiaotong University
Xi'an, China
liuyangzhang@xjtu.edu.cn

Abstract—**In this work we have proposed a novel-shaped metamaterial resonator for terahertz (THz) applications. The proposed resonator consisted of two types of V-shaped metal strip with upward and downward openings and the supportive resin substrate exhibits four high-Q resonance peaks. The surface current and magnetic field distribution at resonance frequencies unveils that incident electromagnetic wave can excite the surface current on the metal strips and form magnetic dipoles, which explains multi-band resonance via the coupling of multi-order magnetic dipoles. Geometric parameters are systemically analyzed to verify the robust THz resonance of proposed structure. It can be expected that our metamaterial structure can be utilized for potential applications including sensing and filtering.**

Keywords—terahertz, metamaterial, magnetic resonance, high-Q factor, coupling effect

I. INTRODUCTION

Metamaterials, artificially manufactured media which are typically made from periodic arrays of subwavelength dimensions, has more extraordinary physical properties than those of conventional materials in nature [1]. With unique engineering permittivity and permeability [2], many unusual physical phenomena such as negative refraction [3], superlenses [4], and optical cloaking [5] have been realized through metamaterial. Many metamaterials-based filters or sensors have been studied in THz range in the past decade, however the Q-factor of these metamaterials are usually less than fifty [6] and the resonant intensity becomes weaker when Q-factor increases. Therefore, it remains a bottleneck to obtain high resonant intensity and high-Q factor at the same time for the metamaterial sensor. For sensing applications, three-dimensional (3D) structure featured with the out-of-plane sensing components has strong capabilities to bond with the sensing medium which allows its high sensitivity for refractive index changes of analytes [7]. And for filtering applications in THz imaging, spectroscopy and other THz communication systems, multi-band filters have more advantages than single-band filter [8]. However, most investigators focused on the filter with single-band resonance [9], which limits practical applications of metamaterial filter.

In this work, a three-dimensional metamaterial resonator with high-Q factor and multi-band resonance is theoretically studied for terahertz applications. To validate the resonant mechanism, magnetic dipoles coupling effects are analyzed via the surface current and magnetic field distribution. Furthermore, the influence of geometric parameters on the resonance effect is also studied.
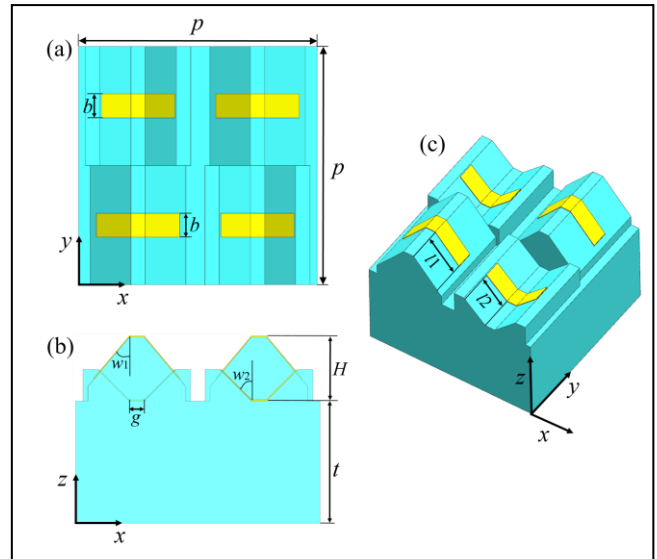
II. DESIGN OF PROPOSED RESONATOR

Fig. 1. Schematic representation of the unit cell of the proposed three dimensional metamaterial resonator. (a) Top view. (b) Front view. (c) Oblique view.

The unit cell of 3D metamaterial resonator is shown in Fig. 1. The resonator is composed of a polymer substrate and four

*Liuyang Zhang is the corresponding author. (e-mail:liuyangzhang@xjtu.edu.cn)

V-shaped metal strips. The geometric dimensions of two V-shaped metal strips along the diagonal position are identical to each other, and the opening of one pair of V-shaped metal strips is upward while the other pair is opposite. Four V-shaped metal strips are evenly distributed in a single unit cell. The convexity and concavity on the polymer substrate can be determined by the bending angles of the upward and downward V-shaped metal strips, respectively. After structural optimization, the optimal geometric size of the structure is initialized with lattice period p as 2000 µm along both x-direction and y-direction. The thickness of the polymer substrate (denoted as t) is 1000 µm. The width of the metal strips (b) is 200 µm and the length of horizontal bars (g) is 120 µm. For the V-shaped metal strips with downward openings, the bend angle (w_1) and length of metal arm (l_1) are 40° and 450 µm, respectively. For the V-shaped metal strips with upward opening, the bend angle (w_2) and length of metal arm (l_2) are 45° and 350 µm, respectively. The relative height (H) between the two kinds of V-shaped strips is 525 µm.

The numerical simulation is carried out with finite element method of the CST Micro Wave Studio in the low terahertz frequency. The material of metal is set as gold and the dielectric substrate is set as resin material with refractive index taken as 1.75 in this work. As showed in the inset of Fig. , the incident electromagnetic is defined as plane wave and the propagating direction is along z-direction. The polarization direction of the electric field is along the x- direction, which is parallel to the length direction of the metal strip in order to induced the magnetic resonance response of the metamaterial. The transmission coefficient of the proposed structure is extracted from the S-parameter which is calculated by the frequency domain solver in the Microwaves&RF/Optical module of the CST software. Periodic boundary conditions are applied along both x- and y-directions of the unit cell structure, with open (adding space) conditions along the z-direction, so the characteristic of periodic arrays in the metamaterials can be simulated correctly.

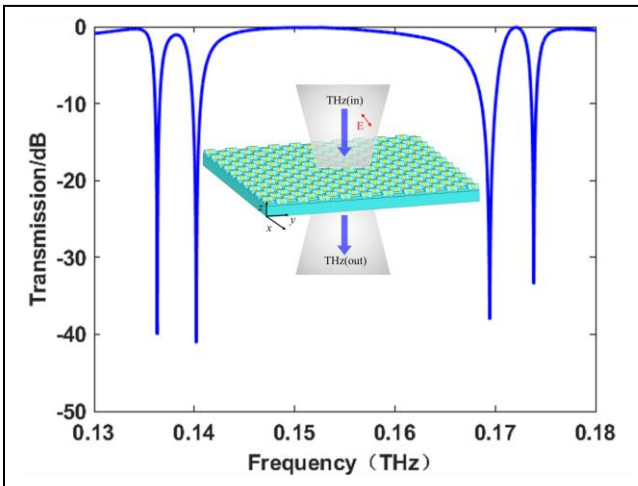## III. SIMULATION RESULTS AND DISSCUSSION

Fig. 2. Transmission spectra of the designed metamaterial resonator, the inset shows the terahertz wave incident and polarization direction.

In the simulation, the complex permittivity of gold in the terahertz region is adequately described with the well-known modified Drude model [10],

$$\varepsilon_{Au} = \varepsilon_{\infty} - \frac{\omega_p^2}{\omega^2 + i\omega\gamma} \tag{1}$$

$$\varepsilon_{Au}' = \varepsilon_{\infty} - \frac{\omega_p^2}{\omega^2 + \gamma^2} \tag{2}$$

$$\varepsilon_{Au}'' = \frac{\omega_p^2 \gamma}{\omega(\omega^2 + \gamma^2)} \tag{3}$$

Formula (1) is the modified Drude model of gold, where ω is the incident angular frequency of the electromagnetic wave, $\varepsilon_{\infty} = 1$ represents the high-frequency bulk dielectric constant of gold, the plasma frequency ω_p is 1.374×10^{16} rad/s and the damping constant γ is 1.225×10^{14} rad/s. Formula (2) and (3) are the real and imaginary part of the dielectric constant respectively. Because of the surface scattering and boundary effect, the damping coefficient of the gold film is three times larger than that of bulk gold. The adaptive meshing refinement is employed to improve the simulation accuracy and computational efficiency.

Fig. shows the transmission spectra of proposed metamaterial resonator as a function of frequency in low terahertz range. The inset picture shows the incident angle (-z direction) and polarization angle (along the x direction) of the incident electromagnetic wave. It can be observed from the transmission spectra that the metamaterial resonator has four resonance bands at frequencies of $f_1 = 0.136$ THz, $f_2 = 0.140$ THz, $f_3 = 0.169$ THz, and $f_4 = 0.174$ THz, respectively. For all these resonance bands, the resonant intensity of their peaks are greater than -30dB, which underlies intensive resonance at these resonant frequencies. Additionally, the quality factor $Q = f_0/\triangle f_{3dB}$ of the resonator is evaluated, where f_0 refers to the resonance frequency, and $\triangle f_{3dB}$ is the -3dB bandwidth of the resonant frequency f_0. All resonance has large quality factor of 96.7, 52.9, 36.8 and 125, respectively. The large Q factor indicates that the loss of the resonance
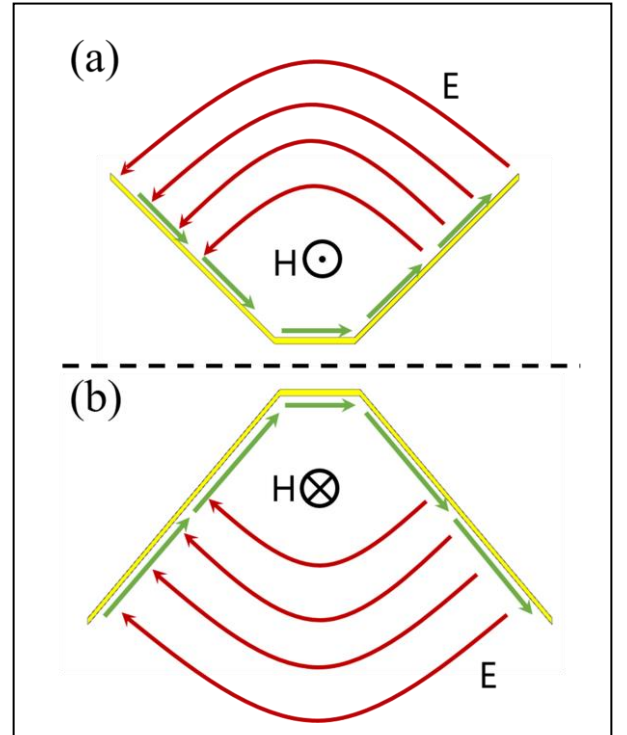

Fig. 3. Magnetic resonance of V-shaped metal strip

system is trivial and the FWHM (full width at half maximum) is small, which is promising for the development of novel electric devices in the sensing field.

As illustrated in Fig. , upon the illumination of normal incident electromagnetic wave, oscillating electric current on the surface of metal strips can be induced by the x-polarization electric field. The induced electric current on the two arms and the horizontal bar can be regarded as the electric dipoles and forms an open circle[11]. In fact, these circulating currents will produce magnetic dipoles along the y-direction and interact strongly with the magnetic field of the incident electromagnetic wave. At resonant frequencies, a strong local electromagnetic enhancement is established between the corresponding metal arms and interact strongly with the incident electromagnetic waves. For the proposed V-shaped metal strip structure, the opening direction and surface current direction determines the direction of the excited magnetic dipoles. For instance, when the surface current direction of the metal strips is the same to the opening direction, the V-shaped

metal strip with upward opening and downward opening will excite the magnetic dipoles in the opposite direction according to the right hand grip rule. On contrary, when the surface current direction is opposite to the opening direction, the excited magnetic dipoles will have the same direction as showed in the Fig. 3 .

To better clarify the physical resonance mechanism of the resonator, the surface current and magnetic field distribution are investigated at each resonance frequency as shown in Fig. 3. At the resonance frequencies f_1 and f_3, the surface current on the upward V-shaped metal strip and the downward V-shaped metal strip are along the same direction as shown in (a) and (c), while at the resonance frequencies f_2 and f_4, the surface current on the upward V-shaped metal strip and the downward V-shaped metal strip are along the opposite direction as shown in (b) and (d). According to the analysis in the previous paragraph, the excited magnetic dipoles have different
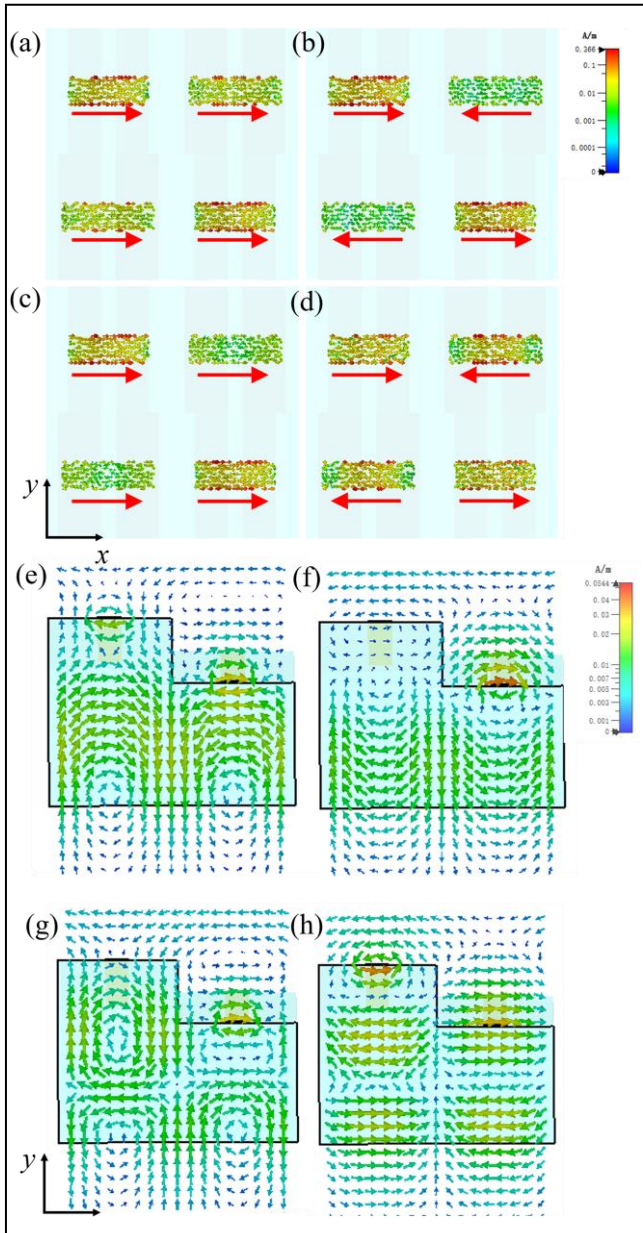

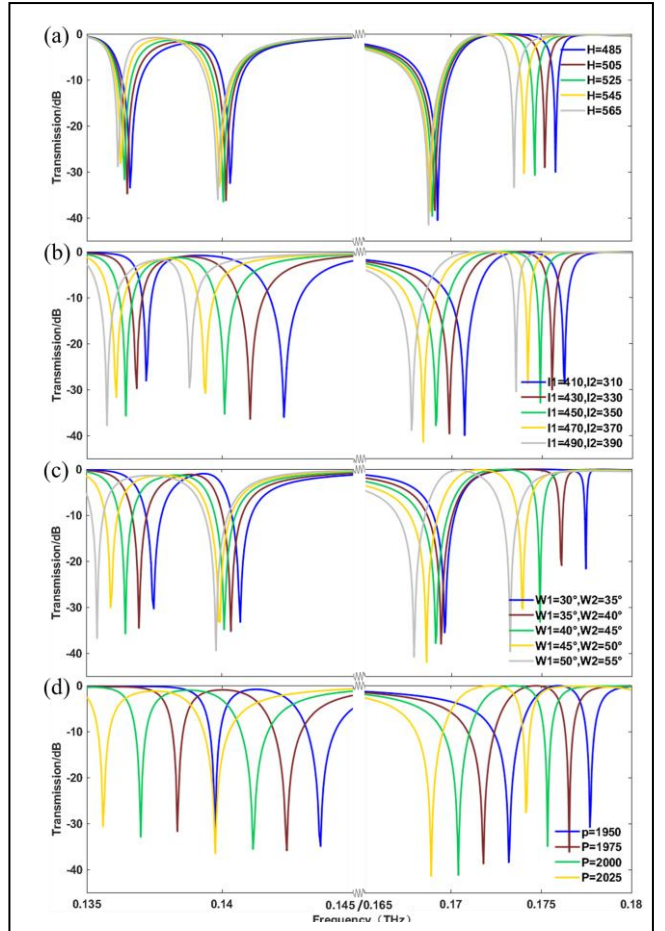Fig. 3. Surface current (a)-(d) and magnetic field (e)-(h) distribution at resonance frequencies.

Fig. 2. Resonance frequencies for different geometric parameters

direction in these circumstance, and the resonance frequencies f_1, f_3 and f_2, f_4 are induced by coupling of the magnetic dipoles with different directions. At resonance frequencies f_2 and f_3, the surface current on the downward V-shaped metal strips is very small, while the upward V-shaped metal strips have high surface current intensity. Apparently, all metal strips have high strength surface current at resonance frequencies f_1 and f_4.

Fig. 3 (e)-(h) shows the magnetic profiles at the resonance frequencies along the y-z plane at $x = 500\ \mu$m, it can be seen that the magnetic field distribution at each resonant frequency agrees well with the surface current distribution. When the

surface current is weak, the magnetic field is also weak. The incident electromagnetic field excites multi-order magnetic dipoles inside the resonator, and the transverse and longitudinal coupling effects between the excited magnetic fields contributes to the appearance of multiple resonant transmission spectra peaks. Low-order magnetic dipoles indicates low frequency resonance (f_1 and f_2) while high-order multiple dipoles coincides with high frequency resonance (f_3 and f_4). Two V-shaped metal strips with opposite openings can excite multi-order magnetic dipoles and these magnetic dipoles coupled with each other result in the multi-order resonance.

To examine the influence of the structure geometric parameters on the resonance characteristics of the metamaterial resonator, the unit cell structure with different geometric parameters as relative height H, length of metal arm l_1 and l_2, bend angle w_1 and w_2 as well as the lattice period p are parametrically studied. Unless otherwise states, b, g and t are fixed in the following simulation setup. As shown in Fig. 2 (a), the resonance peak frequencies have a slight red shift when increasing the relative height H monotonically. As shown in Fig. 2 (b) and (c), the resonance peak frequencies follows similar red-shift with the increment of arm length l and bend angle w, which is attributed to the reason that the projection length of V-shaped metal strips on the x-y surface become larger when l and w increases, and the distance of electron oscillations increases.

In comparison with Fig. 2 (a), (b) and (c), a significant red-shift of the resonance frequencies can be observed when slightly increasing the lattice period p of the resonance structure as shown in Fig. 2 (d), which illustrated that the coupling effect between unit cells has a great effect on the resonance frequency. Interestingly, the resonant intensity is always kept at a high level which exhibits that the three-dimensional resonator is insensitive to the structure change and the resonant band can be extended to other electromagnetic wave frequency range by varying the geometric parameters.

IV. CONCLUSIONS

In this work, we have systematically presented demonstrated multi-band THz response of three-dimensional metamaterial resonator. It shows that the proposed resonator has multi-band high-Q resonance peaks. The surface current and magnetic field distribution indicate that the multi-order magnetic dipoles coupling results in the appearance of resonance frequency peaks. Moreover, by changing the structure size, the resonator can be applied to other frequency bands. The designed metamaterial resonator exhibits stable resonance effect characteristics making it suitable for sensing and filtering technology.

REFERENCES

[1] C. M. Soukoulis, S. Linden, and M. J. S. Wegener, "Negative refractive index at optical wavelengths," vol. 315, no. 5808, pp. 47-49, 2007.

[2] R. A. Shelby, D. R. Smith, and S. Schultz, "Experimental Verification of a Negative Index of Refraction," vol. 292, no. 5514, pp. 77-79, 2001.

[3] X. J. S. C.-p. M. Luo and Astronomy, "Principles of electromagnetic waves in metasurfaces," vol. 58, no. 9, pp. 594201-594201, 2015.

[4] S. Haxha, F. Abdelmalek, F. Ouerghi, M. D. B. Charlton, A. Aggoun, and X. J. S. R. Fang, "Metamaterial Superlenses Operating at Visible Wavelength for Imaging Applications," vol. 8, no. 1, pp. 1-15, 2018.

[5] Z. Jiang, Q. Liang, Z. Li, P. Lv, T. Chen, and D. Li, "Experimental Demonstration of a 3D‐Printed Arched Metasurface Carpet Cloak," *Advanced Optical Materials,* vol. 7, no. 15, 2019.

[6] A. Tavousi, M. A. Mansouribirjandi, M. Ghadrdan, and M. J. P. N. C. Ranjbartorkamani, "Application of photonic crystal ring resonator nonlinear response for full-optical tunable add–drop filtering," vol. 34, no. 1, pp. 131-139, 2017.

[7] Y. Cheng, X. S. Mao, C. Wu, L. Wu, and R. Gong, "Infrared non-planar plasmonic perfect absorber for enhanced sensitive refractive index sensing," *Optical Materials,* vol. 53, pp. 195-200, 2016.

[8] L. Bibbo, K. Khan, Q. Liu, M. Lin, Q. Wang, and Z. J. P. R. Ouyang, "Tunable narrowband antireflection optical filter with a metasurface," vol. 5, no. 5, pp. 500-506, 2017.

[9] I. Alnaib, C. Jansen, N. Born, and M. J. A. P. L. Koch, "Polarization and angle independent terahertz metamaterials with high Q-factors," vol. 98, no. 9, p. 091107, 2011.

[10] M. A. Ordal, R. J. Bell, R. W. Alexander, L. L. Long, and M. R. J. A. O. Querry, "Optical properties of fourteen metals in the infrared and far infrared: Al, Co, Cu, Au, Fe, Pb, Mo, Ni, Pd, Pt, Ag, Ti, V, and W," vol. 24, no. 24, pp. 4493-4499, 1985.

[11] X. Xiong *et al.*, "Polarization-dependent perfect absorbers/reflectors based on a three-dimensional metamaterial," *Physical Review B,* vol. 88, no. 11, 2013.

Modified Generative Adversarial Network for Super-Resolution of Terahertz Image

Zhen Zhang
Xi'an Jiaotong University Shenzhen Academy, Nanshan District, Science and Technology Park
Shenzhen,China
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
zz3119101101@stu.xjtu.edu.cn

Liuyang Zhang *
Xi'an Jiaotong University Shenzhen Academy, Nanshan District, Science and Technology Park
Shenzhen,China
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
liuyangzhang@xjtu.edu.cn

Xuefeng Chen
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
chenxf@xjtu.edu.cn

Yafei Xu
Xi'an Jiaotong University Shenzhen Academy, Nanshan District, Science and Technology Park
Shenzhen,China
State Key Laboratory for Manufacturing Systems Engineering
Xi'an Jiaotong University
Xi'an, China
xyf2492229210@stu.xjtu.edu.cn

Abstract—**Terahertz (THz) images have low spatial resolution, blurring contour features and high background noise owing to the limitation of terahertz (THz) wavelengths and the THz imaging systems. We have proposed a modified Generative Adversarial Network (GAN) for super-resolution (SR) purpose. To fit the THz images, we design a kind of image degradation model to generate low-resolution images with Gaussian blur and white Gaussian noise. We establish a dataset of damage images in the field of non-destructive testing (NDT) for training and testing. The experimental results on THz images demonstrate that the improved GAN model can improve the quality of THz images effectively. Our method can be beneficial to improve the accuracy of THz NDT with low resolution.**

Keywords—THz image, super-resolution, degradation model, deep learning

I. Introduction

Terahertz (THz) wave [1] is a kind of electromagnetic radiation with strong penetrability and harmless photon energy, which is widely used in the field of non-destructive testing (NDT) [2,3]. However, due to the limitation of the THz wavelength and the agonizingly slow development of THz imaging technology, it encounters severe problems in THz images include low spatial resolution, blurring contour features and high background noise [4]. It is obligatory to upgrade or optimize the THz imaging system and the algorithm to overcome these drawbacks. Actually, Huang et al. [5] have tried to design a kind of metallic grating to realize the super-resolution imaging in THz frequency, which is considered as a landmark of the high resolution THz imaging system. However the grating-assisted imaging system is tightly constrained to its expensive hardware and poor effect. It is extremely necessary to come up with appropriative super resolution (SR) algorithms of THz image. Ding et al. [6] have adopted Lucy–Richardson method to recover images with relatively low-resolution obtained from a THz reflective imaging system. However, it's difficult to guarantee the convergence via the Lucy–Richardson algorithm when the images have low signal-to-noise ratio (SNR), and it's possible to amplify noises. Kiarash et al. [7] have proposed a method combined THz point spread function (PSF) and deconvolution for THz image restoration. Hu et al. [8] have presented an aggregate wavelet-predominant algorithm for THz security images which focused more on the image sharpness. These methods improve the resolution of THz images to some extent, but there is still larger improvement space. Besides interpolation-based and reconstruction-based methods, learning-based [9] method is also widely used in the field of THz with enhanced resolution.

CNN is the most widely used learning-based method for THz images by mapping low-resolution (LR) image to high-resolution (HR) image. Long et al. [10] have demonstrated that the deep Convolutional Neural Network (CNN) model is effective and robust for the super resolution of THz image. But there remains some shortcomings in the recovery process of high-frequency details via CNN. In 2016, Generative Adversarial Network (GAN) based SR [11] was proposed by Ledig et al. by using generator G to generate SR images though LR images, and the discriminator D to determine whether the images are generated by G or the original images. However, its training process is based on the natural image datasets so that it is not suitable for THz images. At the same time, there are severe training difficulties in GAN such as overfitting and mode collapse.

In this work, by taking full advantages of GAN, we propose an modified GAN model for super resolution of THz image. Multiple damage images from NDT are used to establish the dataset for experiments. Pre-training and fine tuning are used to optimize the training process and address the impact of the dataset size. Experimental results indicate that the proposed model is an effective super resolution method for LR THz images.

* Liuyang Zhang is the corresponding author. (e-mail: liuyangzhang@xjtu.edu.cn).
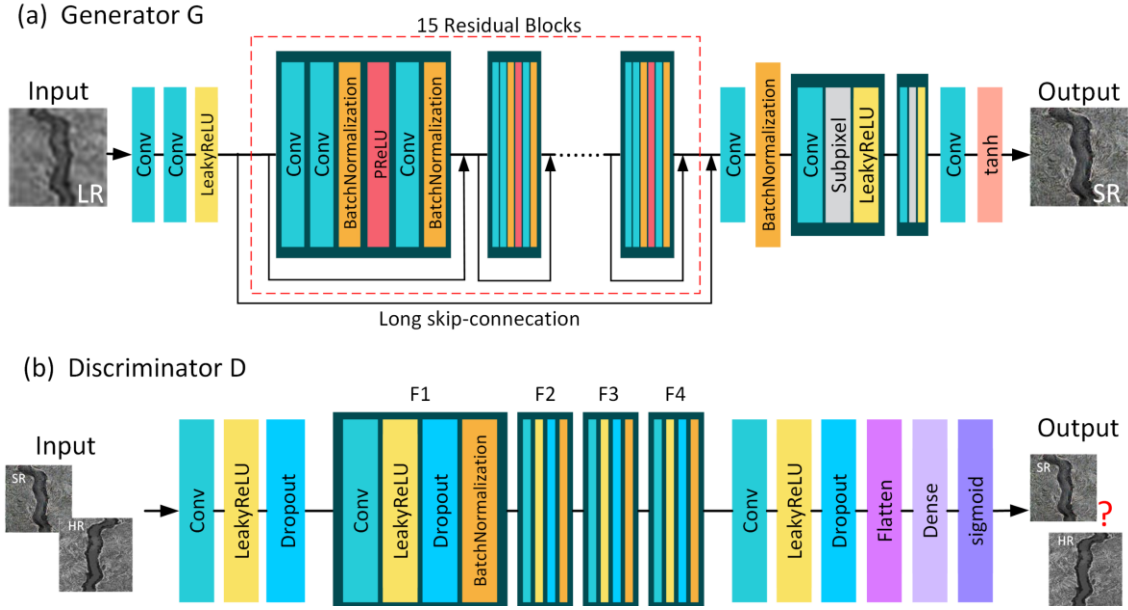
Fig. 1. Architecture of the improved Generative Adversarial Network: (a) is the generative network, (b) is the discriminator network which has 4 different blocks.

II. PROPOSED METHOD

A. Degradation Model

In order to learn the relationship between LR images and HR images, it's necessary to design the degradation model. LR images with THz-like properties are obtained by the degradation model. Jepsen et al. [12] demonstrate that the background blur and noise in THz image follows approximately Gaussian distribution. A appropriative degradation model for THz image can be written as:

$$y = x \otimes b + n \qquad (1)$$

where x is the HR image, b denotes the Gaussian blur kernel, $\otimes$ is the convolution operation and n is the white Gaussian noise to simulate the THz background noise. The Gaussian blur kernel is generated as follow:

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \qquad (2)$$

where σ denotes the Gaussian radius in image processing, σ means width of the sliding window. In order to obtain the LR image similar to THz image and reduce the training burden of the network, σ is set as 3.0. The standard deviation of white Gaussian noise is set as $\hat{\sigma} = 1.0$.

B. Modified GAN model

The modified GAN architecture is illustrated in Fig.1. The generative network in Fig.1 (a) has 15 residual blocks. Two 9×9 convolution layers and a activation function LeakyReLU are employed at the first three layers to extract shallow features. LeakyReLU is a variant of ReLU which can overcome the "dead problem" in ReLU in the training process. To extract more features from THz images and make the generator G perform better, a deep network structure is build. But with the deepening of the network structure, gradient vanishing and gradient explosion will occur [13]. Previous researches show that the residual learning framework could address these issues and the skip connection between layers could train a deep network

effectively. Therefore, the residual learning framework is added to improve the performance of SRGAN. It contains 15 residual blocks and a long skip connection, each residual block also has a short skip connection. Following the residual learning framework, a 9×9 convolution layer and a batch normalization are placed to reduce the instability of parameter distribution and enhance the generalization ability of the network. A kind of sub-pixel block consisting of the sub-pixel convolution is designed to enlarge the feature map, which could avoid artificial factors. Activation function tanh is employed at the last layer to normalize the range of elements in training image matrix between -1 and 1. Except the 9×9 filters mentioned above, all other convolution layers use 3×3 filters.
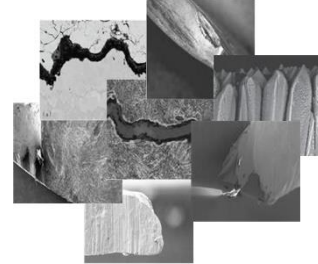

Fig. 2. Dataset of damage images in the field of NDT.

It should be noticed that serious shortcoming of the GAN is that generator G and discriminator D are difficult to converge at the same time. To find a Nash equilibrium [14] between the two networks and achieve an optimal convergence, a simple network structure as the discriminator D is designed. The discriminator D in Fig.1 (b) is used to solve the binary classification problem between HR images and SR images. A 3×3 convolution layer, a activation function LeakyReLU and a dropout layer are employed at the first three layers to extract features. The dropout layer is used to reduce training parameters and overcome the overfitting. Following the first part, four blocks with different channels are designed. Following the dropout layer, a batch normalization layer is added in each
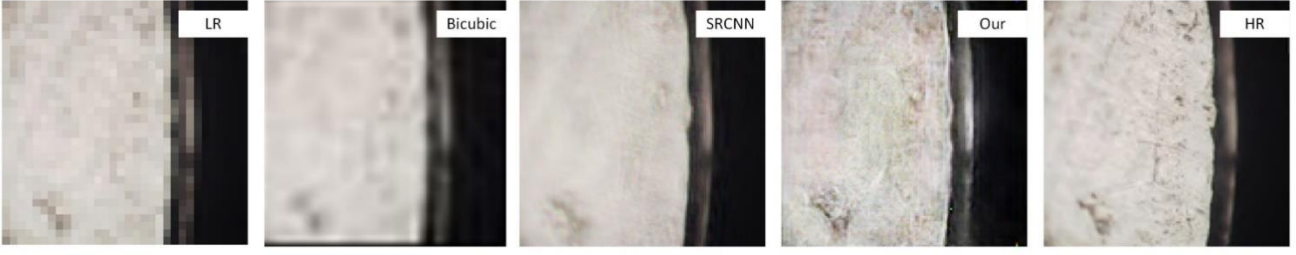
Fig. 3.   Damage image "test4" super resolution results of our model and other methods.
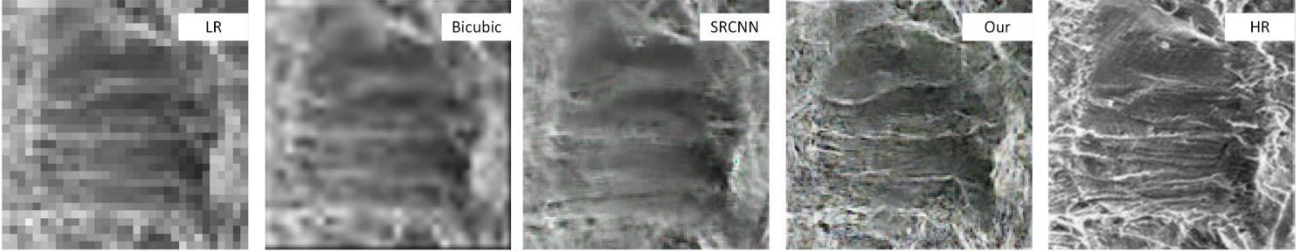

Fig. 4.   Damage image "test5" super resolution results of our model and other methods.
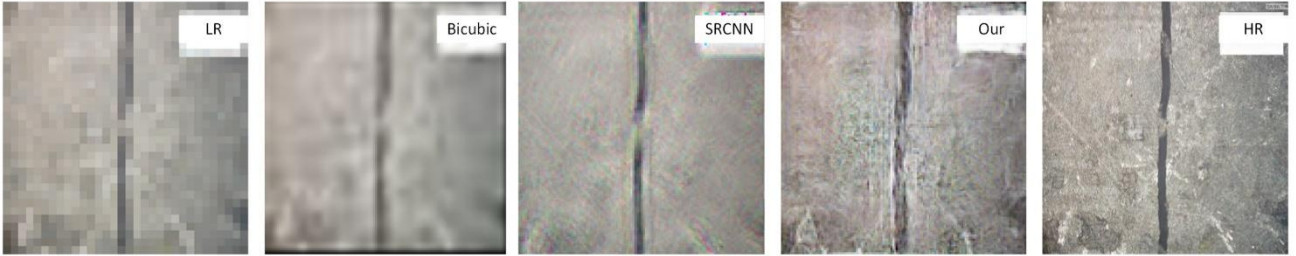

Fig. 5.   Damage image "test6" super resolution results of our model and other methods.

block. The batch normalization layer is used to build different batches for real data and fake data. The final part with a dense layer and a sigmoid layer is used to limit the range of the output between 0 and 1. All convolution layers use 3×3 filters.

The THz-like LR images are imported into the generator. Based on the relationship learned by the GAN model, SR images are obtained. Then the SR images and corresponding HR images are imported into the discriminator. Discriminator judges whether the generated images are original images. Target images are obtained via the game between the generator and the discriminator.

III. EXPERIMENTS

A. Parameters Setting

During the training processing, when the learning rate of generator G or discriminator D is changed, the other's is retarded until reach a stable state finally. Adam (15) is an extension of stochastic gradient descent algorithm, which is widely used in the field of image super-resolution. For SRGAN, both generator G and discriminator D use Adam with $\beta = 0.9$ as optimizers. In order to achieve a stable convergence state quickly, different optimizers for the two sub-networks are used.

For the generator G we use Adam with $\beta_1 = 0.5$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$ and the initial learning rate lr = 0.0002. For the discriminator D we use RMSprop with initial learning rate as lr = 0.00002 . Both two learning rates decay every 100 epochs. Additionally, we adopt weight initialization in all convolution layers with weight decay as 0.0001.

The Tensorflow 10.0 + keras 2.0 is employed to do our experiments for the modified GAN model. Training takes 5 days on the RTX 2080Ti GPU.

A. Datasets and Training Setup

For general super-resolution, natural images are widely used as training or testing datasets. However, THz images are quite different from natural images. For THz image super-resolution, we had to pay more attention to the contour features and edge features. Therefore, a dataset containing 1771 mechanical structural defects in Fig.2 is created by collecting the data from the published paper. 80% of the dataset are used as the training dataset and others are used as the testing dataset. Finally, 50 THz images are used to test the model's performance.

For image super-resolution, the size of the dataset will determine the quality of network training. If the size is not large enough for a complex model, it's easy to be trapped into overfitting. Thus we train the proposed model over a large dataset included 11100 natural images firstly to get a pre-training model. Then the pre-training model is retrained on the home-established dataset to adjust the hyperparameters over the dataset. To reduce the calculation of the model, all LR image's size in the datasets is 32×32. The batch size of the training datasets is 64. A mini batch size is set to 28 to make the discriminator D and the combined model update weights respectively.

B. Experiment Results

To demonstrate the robustness of our proposed GAN model, the performances of Bicubic, SRCNN and our model on the testing dataset are systemically evaluated, respectively. The LR images and restored images are shown

in Fig.3, Fig.4 and Fig.5, respectively. Obviously, our model has a better performance than other methods.

In Fig.3, the SR images obtained via Bicubic and SRCNN don't have obvious texture details. Only the edge contour can be restored to a certain extent. Our model not only restores texture details, but also generate the sharp contour. In the field of non-destructive testing, getting more high-frequency details means more accurate judgment of the damage. It can be seen in Fig.4 and Fig.5, for complex image details, our model has better performance.

In the TABLE I, performances are further quantitively verified by peak signal-to-noise ratio (PSNR). PNSR is an important evaluation index to measure the quality of the reconstructed images. Generally, there are less distortions of the reconstructed image when its value of PSNR is larger. As is shown in TABLE I, our model has better performances in the three images mentioned above

To demonstrate the ability of our model, 50 random THz images are used to test the model. In Fig.6, it can be seen that our model performs better than other methods, where the contours of the THz image is clearly restored.
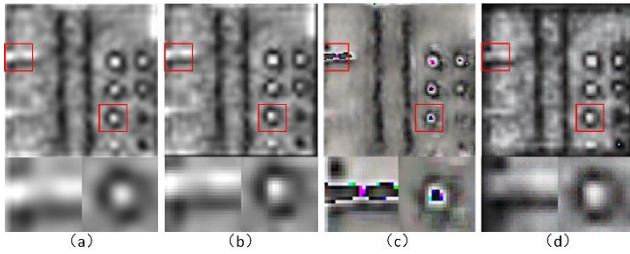
Fig. 6. Experiment results on the THz image. (a) THz image, (b) Bicubic, (c) SRCNN and (d) our model.

TABLE I.   PSNR RESULTS OF DIFFERENT METHODS ON THE TESTING DATASET

Test images	PSNR		
	Bicubic	*SRCNN*	*Ours*
test4	24.38	26.95	28.47
test5	20.41	23.66	26.92
test6	21.66	22.48	22.96

IV. CONCLUSION

In this work, we have proposed a modified GAN model for super resolution of THz image. The improvement of the generator G and the discriminator D make the overall framework fit our datasets. Instead of using down-sampling, we design a degradation model with Gaussian blur and white Gaussian noise for THz image. These two parts correspond to the blurring contour features and high background noise in THz images respectively. We have a

pre-training on a huge dataset to overcome the overfitting and then adjust the hyperparameters on our datasets to get the optimal model. The results demonstrate that the proposed model performs better than other method on the quality of LR THz images.

REFERENCES

[1] B. B. Hu, and M. C. Nuss, "Imaging with terahertz waves," Opt. Lett., vol.20, pp.1716–1718, 1995.

[2] P. U. Jepsen, D. G. Cooke, and Koch, M., "Terahertz spectroscopy and imaging – Modern techniques and applications," Laser & Photonics Reviews, vol.5, pp.124–166, 2011.

[3] C. Jansen, et al., "Terahertz imaging: applications and perspectives," Applied Optics, vol.49, pp.48–57, 2010.

[4] L. M. Xu, W. H. Fan, and J. Liu, "High-resolution reconstruction for terahertz imaging," Applied Optics, vol.53, pp.7891–7897, 2014.

[5] T. J. Huang, H. H. Tang, Y. Tan, L. Li, and P. K. Liu, "Terahertz super-resolution imaging based on subwavelength metallic grating," IEEE Transactions on Antennas and Propagation, vol.67, pp.3358-3365, 2019.

[6] D. Sheng-Hui, L. Qi, Y. Rui, and W. Qi, "High-resolution terahertz reflective imaging and image restoration," Appl. Opt., vol. 49, no. 36, pp. 6834--6839, 2010.

[7] K. Ahi and M. Anwar, "Developing terahertz imaging equation and enhancement of the resolution of terahertz images using deconvolution," in Terahertz Physics, Devices, and Systems X: Advanced Applications in Industry and Defense, vol. 9856, M. F. Anwar, T. W. Crowe, and T. Manzur, Eds. (Proceedings of SPIE. Bellingham: Spie-Int Soc Optical Engineering, 2016.

[8] M. Hu et al., "A wavelet-predominant algorithm can evaluate quality of THz security image and identify its usability," IEEE Transactions on Broadcasting, vol. 66, no. 1, pp. 140-152, 2020.

[9] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," Proc. Eur. Conf. Comput. Vis. Springer, 2014.

[10] Z. Y. Long et al., "Terahertz image super-resolution based on a deep convolutional neural network," Applied Optics, 2019.

[11] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," 2016.

[12] P. Jepsen, R. Jacobsen, and S. Keiding, "Generation and detection of terahertz pulses from biased semiconductor antennas," . Opt. Soc. Am. B, vol.13, pp. 2424-2436, 1996.

[13] Z. Lu et al., "Fast single image super-resolution via dilated residual networks," IEEE Access, pp. 1-1, 2018.

[14] T. Salimans et al., "Improved techniques for training GANs," pp. 2234--2242, 2016.

[15] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," Computer Ence, 2014.

A Thermal-imaging-based Method for 2D Electric Current Distribution Measurement

Chao Ren
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
rxc7250@126.com

Lijian Zhu
Shanghai Space Propulsion Technology
Research Institute
Shanghai, China
zljnpu@163.com

Libing Bai*
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
libing.bai@uestc.edu.cn

Lulu Tian
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
tl.110119@163.com

Jie Zhang
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
zhj06_19@uestc.edu.cn

Zhen Liu
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
scdliu@uestc.edu.cn

Abstract—**Measurement of electric current distribution is important for medicine, semiconductor, and structure integrity evaluation in nondestructive testing (NDT). Several electric measurement methods are developed, such as electrode array, coil and magnetic sensor. In this paper, the infrared (IR)-imaging-based method is proposed to capture the 2D electric current distribution. In this method, the current magnitude matrix is obtained by utilizing the relationship between the thermal field and the electric field. Furthermore, an iteration algorithm is proposed to reconstruct the direction of the current depending on the divergence theorem of passive field. Then the 2D electric current distribution can be visualized with high resolution and the results would verify the efficiency of the proposed method.**

Keywords—*2D electric current distribution, nondestructive testing (NDT), infrared (IR)-imaging, divergence theorem, current direction reconstruction*

I. INTRODUCTION

Electric current distribution measurement attracts more and more attentions because of its widely applications. For example, in medical examination, measuring the changes of current distribution caused by lesions of organ in animal or human body which is helpful in locating the lesions [1]. In semiconductors design and applications, electric current distribution is extended to predict the failure behavior of the devices [2] [3]. In order to evaluate the integrity of material structure, eddy current testing method is developed[4]. Furthermore, electric current distribution measurement also play an important role at the cutting-edge of 2D nanodevices (graphene [5], silicene [6], h-BN [7], MoS2 [8]) development [9].

In recent years, several electric current distribution measurement methods are developed. In [10] and [11], Electrode array is used to map the current through potential drop，which indicates the electrical conductivity variation caused by lesion organ. Scanning probe microscope (SPM) is applied to image cyclotron orbits of electrons in graphene [12]. Scanning superconducting quantum interference device (SQUID) [13] and atomic-sized quantum sensors array [14] are applied to map current distribution through magnetic field measurement. All of these methods map current distribution from intermediate physical quantities (voltage, magnetic field) by reconstruction algorithm. In lesion organ image acquisition, the conductivity is obtained by establishing the nonlinear function between voltage and conductivity and finding the minimum value of the least squares problem, which is constructed by the difference between the actual voltage value and the calculated voltage value [11]. When it comes to current density reconstruction, The relationship between the current density and the z-component of the magnetic field is employed through the Biot–Savart law. [13] [14]. Thus, high performance sensors and reconstruction algorithm is the key for current distribution measurement.

However, most of the above-mentioned methods use big size sensors which makes it impossible to capture a high-spatial-resolution image of current distribution. Even SPM and SQUID reduce the size of sensors, they are very expensive and complicated. Therefore, there is a definitely demand to develop an electric current distribution measurement technique with non-contact, high spatial resolution for in-situ application.

Infrared thermography (IR) is an important remote detection method with high spatial-resolution and thermal sensitivity, which has rapidly attracted a lot of attentions in many emerging industrial fields including microelectronics, renewable and sustainable energy (RSE) system, and electric power [15], etc. Wu et al. [16] used IR to detect obvious thermal contrast between the defect-free and defective regions in IGBT, which was influenced by magnetization current intensity and magnetization direction. In [17], He et al. used IR to detect the high temperature areas around the defect owing to resistive heating from the induced electric current. In [18], Zhao et al. used IR to capture the thermal response about local overheating area on insulator, which is caused by high current produced by a short circuit or improper electrical loading conditions. However, all of these thermal-imaging-based methods are qualitative detection methods, which only detect high temperature areas due to current accumulation with no consideration about electric current quantization issue. Since all of electronic parts and components generate heat when working, the higher of electric current, the higher of temperature. Therefore, it is possible to measure electric current distribution by thermography. This is also the purpose of this work.

When high-frequency eddy current excitation source is applied to the specimen, the heat distribution inside the specimen is uneven, and there will be three heat transfer modes in the specimen: heat convection, heat radiation and heat conduction. In order to establish relationship between the thermal field and the electric field, assuming that impact of heat convection and radiation are negligible. Thus, according to Fourier's law, the heat conduction of test specimen is obtained as:

$$\rho C \frac{\partial T}{\partial t} - \nabla(k\nabla T) = P_w \qquad (1)$$

where k is the coefficient of thermal conductivity, P_w is the thermal power. Based on equation (1), it is gotten that:

$$\frac{\partial T}{\partial t} = \frac{I^2 R + \nabla(k\nabla T)}{\rho C} \qquad (2)$$

where $\nabla(k\nabla T)$ is represented the temperature gradient. When the specimen thickness is very small and the heating time is very short, the temperature gradient will be very small. Thus, the effect of temperature gradient can be ignored, and the equation (2) can be expressed as:

$$\frac{\partial T}{\partial t} = \frac{I^2 R}{\rho C} \qquad (3)$$

Considering that ρ, C, and σ are constant, it is concluded that the slope of the temperature-time curve $\frac{\partial T}{\partial t}$ is proportional to the current amplitude matrix $I(x,y)$. Therefore, the current intensity I at each point (x,y) from the infrared image can be obtained by equation (3).

In order to obtain the direction of the induced current, the current amplitude matrix $I(x,y)$ is divided into the vertical component $VD(x,y)$ and the horizontal component $HD(x,y)$. Where $VD(x,y)$ and $HD(x,y)$ satisfy the conditions as:

$$VD(x,y)^2 + HD(x,y)^2 = I(x,y)^2 \qquad (4)$$

Divergence is the flux of a vector F emanating from a point or a closed surface, which can be explained as:

$$divF = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} + \frac{\partial F}{\partial z} \qquad (5)$$

When $divF = 0$, it represents the passive field. the total amount of eddy current entering the surface is equal to the current flowing out of the surface. Based on the divergence of eddy current is 0, the four adjacent pixels are treated as a closed surface shown in Fig.1. Where $a_i (i=1\sim9)$ represents 9 pixels and the red box is the closed surface, the sum of all the eddy current passing through the closed surface should be theoretically equal to 0. The vertical component and horizontal component of point a_i are supposed as a_i^1

and a_i^2. Provided that the right direction and the vertical direction are positive direction, and if the actual induced current direction is as shown in Fig.1, the total amount of induced current on the closed surface formed in the red closed box is:

$$sum = \left| -a_1^1 - a_1^2 - a_2^1 + a_2^2 + a_3^1 - a_3^2 + a_4^1 + a_4^2 \right| \qquad (6)$$
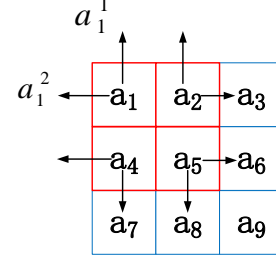

Fig.1 Induced current direction analysis based on divergence is equal to zero of passive field.

As shown in Fig.2, there are four closed surface formed by 9 pixels (red closed box, green closed box, blue closed box, purple closed box). The center point a_5 is supposed as target point which is also the intersection point of the four closed surfaces. These four closed surfaces form a 3×3 square.
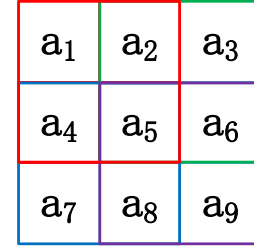

Fig.2. The target point a_5 participates in the formation of four closed surfaces

When each deviation value sum from these four closed surfaces are calculated, the total amount $\sum sum$ can be given in:

$$\sum sum = \left| \sum_{i=1}^{4}\sum_{j=1}^{2} \pm a_i^j \right| + \left| \sum_{j=1}^{2}(a_2^j + a_3^j + a_5^j + a_6^j) \right|$$
$$+ \left| \sum_{j=1}^{2}(a_4^j + a_5^j + a_7^j + a_8^j) \right| \qquad (7)$$
$$+ \left| \sum_{j=1}^{2}(a_4^j + a_5^j + a_7^j + a_8^j) \right|$$

According to the equation (7), the following iterative algorithm can be proposed: Except the points on the edge of the image, each pixel in the infrared image is involved in forming a closed surface contained four pixels with using square cell, and can form four closed surfaces. For any four closed surfaces of one target point in the thermal image, the value of the $\sum sum$ should be equal to 0. the objective function can be expressed as:

$$\min \sum sum \qquad (8)$$

When the vertical component of the target point changes, the horizontal component of the target point will also change due to the constraint of equation (17). If the value of $\sum sum1$ after the change of the vertical component is less than the value of $\sum sum0$ before the change, it indicates that the change of the target point is valid, and the result of the change will be retained.

In the initial case, the value of all elements in the horizontal component $HD(x, y)$ is set to zero, and the value of vertical component $VD(x, y)$ is equal to induced current amplitude matrix $I(x, y)$. Thus, the initial direction is vertical. After calculated the initial deviation value $\sum sum0$. Suppose the step size is Δ which is set to control the convergence speed and accuracy, the specific update strategy is:

$$\begin{cases} VD1(x, y) = VD(x, y) - \Delta \\ HD1(x, y) = \sqrt{I(x, y)^2 - VD1(x, y)^2} \end{cases} \quad (9)$$

$$\begin{cases} VD1(x, y) = VD(x, y) - \Delta \\ HD2(x, y) = -\sqrt{I(x, y)^2 - VD1(x, y)^2} \end{cases} \quad (10)$$

After the vertical component $VD(x, y)$ reduced by Δ, the horizontal component $HD(x, y)$ and vertical component $VD(x, y)$ of the target point a_5 are replaced by $HD1(x, y)$, $VD1(x, y)$ in equation (18) and the deviation value $\sum sum1$ is calculated. In the same way, the deviation value $\sum sum2$ is calculated by $HD2(x, y)$ and the $VD1(x, y)$ in equation (19). Furthermore, the vertical component $VD(x, y)$ increased by Δ is also considered:

$$\begin{cases} VD2(x, y) = VD(x, y) + \Delta \\ HD3(x, y) = \sqrt{I(x, y)^2 - VD2(x, y)^2} \end{cases} \quad (11)$$

$$\begin{cases} VD2(x, y) = VD(x, y) + \Delta \\ HD4(x, y) = -\sqrt{I(x, y)^2 - VD2(x, y)^2} \end{cases} \quad (12)$$

where the horizontal component $HD(x, y)$ and vertical component $VD(x, y)$ are replace by the $HD3(x, y)$, $VD2(x, y)$, and the $HD4(x, y)$ and the $VD2(x, y)$. The deviation value $\sum sum3$, $\sum sum4$ are also calculated. Compared the five initial deviation value $\sum sum0 \sim \sum sum4$, the horizontal component and vertical component of the smallest deviation value are selected to replace the initial horizontal component $HD(x, y)$ and the vertical component $VD(x, y)$. It is worth noting that if the vertical component $VD(x, y)$ is equal to $I(x, y)$ or zero, the step size doesn't need to increase or decrease. In this case, only three deviation value are necessary to compare. The above iterative process is carried out for every point except the edge pixel, until it is impossible to change the vertical

component and horizontal component of any pixel. It is indicating that the horizontal component and vertical component of each pixel are the optimal solution, which is satisfied passive field divergence is equal to zero. Afterwards the direction of induced current can be gained by using quiver function of MATLAB. Thus, it is concluded that 2D current distribution measurement method can be achieved by using proposed iterative algorithm.

III. EXPERIMENT STUDIES

A. Sample Fabrication

In order to reconstruct current direction, steel was used to made the sample for verification, because it is one of the most commonly material in railway, oil pipeline, ship, bridge and so on, and has a strong demand of integrity evaluation. Furthermore, the properties of the steel causing its heating style is surface heating, which make the diffusion effect of steel and the temperature gradient are very low. The lateral dimension of steel is 100×45 mm, and the thicknesses is 0.24 mm. There is one man-made slot with lateral size of 10×2 mm at the center of the specimen.

B. Experimental System

In the current direction reconstruction experiment, both induction coil and IR camera were placed on the same side, as shown in Fig.3 (A). When an excitation coil with a certain frequency AC current was close to the steel sample, it produced eddy current with the same frequency in the sample. Due to the skin depth of steel is very small, the induced current is concentrated on the surface. As shown in Fig. (B), The excitation coil was made of 6.35-mm high-conductivity hollow copper tubing. 256kHz tone burst eddy current (0.1s) was introduced into the steel sample to stimulate temperature variation. The infrared radiation was recorded by using a infrared camera (FLIR SC7500) with frame rate 383Hz. At 383Hz frame rate, the resolution of infrared camera is 320×256. The unit of radiation is digital level(DL).
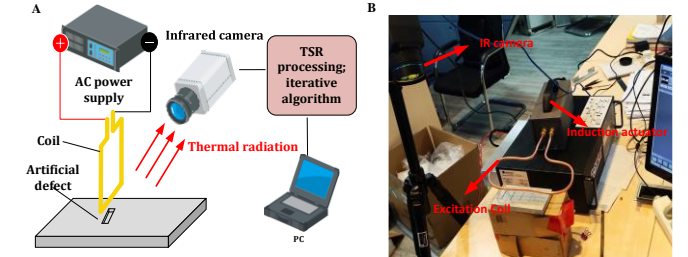

Fig.3 (A) The experimental schematic diagram of current direction reconstruction method. (B) The equipment of current direction reconstruction method

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In terms of thermal sequence data processing, original temperature signal is not smooth enough, derivative $\dfrac{\partial T(t)}{\partial t}$ at the sharp bending point is discontinuous and even abrupt change, which will seriously affect the calculation of induced current amplitude I. Thus, temperature signal reconstruction (TSR) is used here to smooth the original temperature signal and enhance image spatial resolution. Compared to

the difference method, TSR can compensate for the

sampling frequency which is lower than AC current frequency, and obtain the transient rate of temperature change $T(t)$.

Through the TSR processing, the infrared image sequence can be transformed into to a reconstructed image with the same size as the original image, and the effect of environmental noise on infrared image is significantly suppressed. Thus, the smooth temperature-time response $T(t)$ can be obtained by n degree polynomial fitting as:

$$T(t) = e^{a_0 + a_1 \ln(t) + a_2 \ln(t)^2 + ... + a_n \ln(t)^n} \tag{13}$$

where a_n is represent the coefficient after logarithmic domain fitting. Based on taking derivation to the curve $T(t)$ of each point in the thermal image, the slope of the temperature-time curve $\frac{\partial T(t)}{\partial t}$ can be obtained, which is linear to the square of current amplitude I.
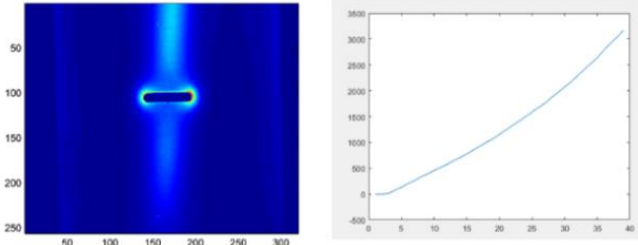
Fig.4 Infrared image of frame 39 and temperature-time response curve of the (98,142).
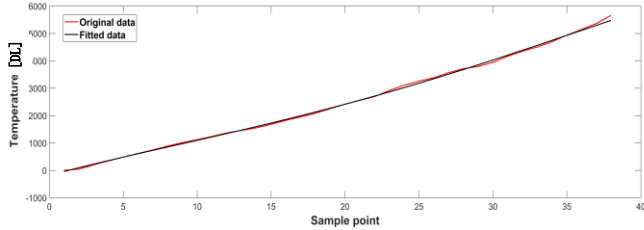
Fig.5 the polynomial fitting result of of point (98,142) in the thermal image.

It is already known that the value of heating time t should be as small as possible to reduce the influence of the diffusion effect of heat. Therefore, the heating time was set to 0.1s. In the direction reconstruction method of current, the steel sample with known man-made slot is heated by the Easyheat 224 firstly. 256kHz tone burst eddy current is introduce into the steel to stimulate temperature variation. The infrared radiation is recorded using an infrared camera with frame rate 383Hz, then the first 39 frames infrared thermal plot (heating period) at each point were smooth processed by using N order polynomial fitting function. The thermal image of frame 39 and the entire thermal-time response curve of the point with coordinates of (98,142) is shown in Fig.4. Polynomial fitting performed on this point of the infrared thermal sequence, and the fitting result is shown in Fig.5.

After calculating the $\frac{\partial T(t)}{\partial t}$ from the fitted polynomial of each pixel in the infrared thermal image, the second frame

time of the arithmetic square root of the derivative value $\frac{\partial T(t)}{\partial t}$ was selected as a point of eddy current amplitude values for analysis in order to reduce the impact of diffusion effect of heat. After normalization, the values of eddy current amplitude at each point $I(x, y)$ obtained by the above processing were obtained, as shown in Fig.6.
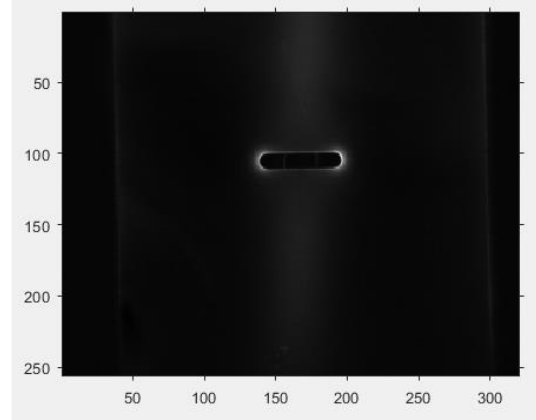

Fig. 6 The diagram of Eddy current density.

As mentioned before, the sum of the eddy current passing through the any closed surface is equal to 0 in the passive field at any given instant. Using this law and physical properties of eddy current field as constraint conditions, the horizontal component and vertical component of eddy current are obtained by using iterative algorithm. The above iterative algorithm was used to process the eddy current density matrix, and the step length was set as 0.05.

The vector images was drawn with the horizontal and vertical components to obtain the current distribution, so as to realize imaging the 2D current distribution. The reconstruction result of the current direction is shown in Fig.7 (A). The current flow around the defect edge was also observed by the local amplification figure in Fig.7 (B), and the length of a vector arrow reflect the amplitude of the eddy current. Thus, the amplitude of eddy current at edge is bigger than other area due to the accumulation of the current at slot edge. But in the area without man-made slot, eddy current density is sparse and the flows are all in the same direction. These features are consistent with the presence of the existing research results about pulsed eddy current thermal imaging, which likely occurred 'hot spot' during the transfer of heat around the defect.

In all, the iterative algorithm proposed in this paper makes full use of the physical meaning of infrared thermal image and can better mine the information in thermal response sequence. A non-invasive current distribution measurement method was achieved. Our approach provides a unique way to capture the 2D electric current distribution based on infrared (IR)-imaging with pixel level spatial resolution, which can be useful for the long-distance and non-invasive measurement method that is hard to be achieved by traditional methods.
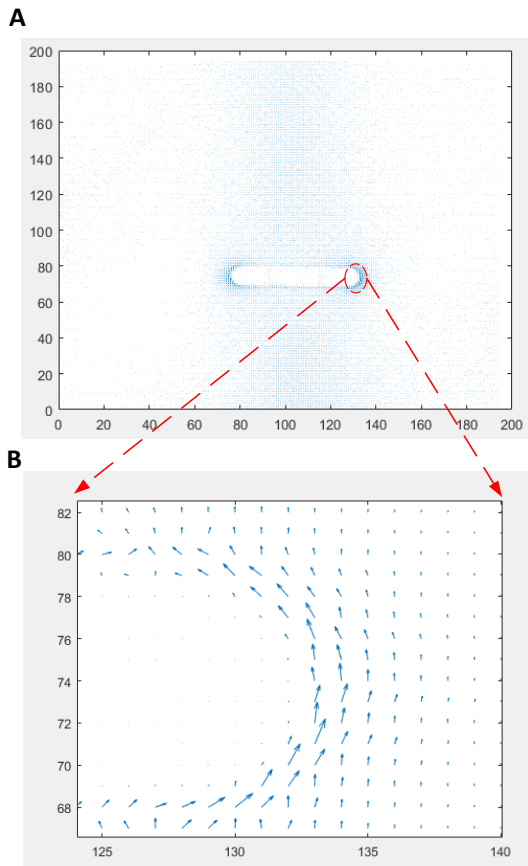
A

B

Fig.7 The reconstruction result of the 2D current distribution. (A) The 2D current reconstruction result including intensity and direction. (B) the local amplification figure of current reconstruction result.

V. CONCLUSION

In this paper, a novel method, which has the ability to map the current amplitude and direction, with advantage of long-distance and non-contact detection, provides an avenue for high-resolution measurement of electric current distribution. This work analyzed the interaction mechanism among temperature variation and electric current amplitude and the characteristics of passive fields. In consideration of IR images are really reflecting the whole current distributions and surface conditions, these signals should be directly correlated with physical-mechanical properties of the metal. The experimental results demonstrate that that the temperature change is linearly proportional to the amplitude of AC current, which allows quantitative measurements for current by thermal signal. Then, based on the divergence theory of passive field, the TSR and iteration algorithm are used to reconstruct the direction of the eddy current. Thus, a method for 2D current detection has been realized. In future work, the proposed approach is also applicable to image the current under strict insulation or multiple points monitoring requirement condition, which will generate significant impact on the development of current measurement. Hence, the current detection methods reported here could become a ubiquitous investigation tools for current distribution measurement in the coming years.

REFERENCES

[1] T. J. Yorkey, J. G. Webster, and W. J. Tompkins, "Comparing Reconstruction Algorithms for Electrical Impedance Tomography," IEEE Trans.biomed.eng, vol. BME-34, no. 11, pp. 843-852, 2007.

[2] G. Qin, C. Lu, X. Zhang, and M. Zhao, "Electric Current Dependent Fracture in GaN Piezoelectric Semiconductor Ceramics," Materials, vol. 11, no. 10, pp. 2000, 2018.

[3] J. Sladek, V. Sladek, H. H.-H. Lu, and D. L. Young, "The FEM analysis of FGM piezoelectric semiconductor problems," Composite Structures, vol. 163, no. MAR., pp. 13-20, 2017.

[4] J. García-Martín, J. Gómez-Gil, and E. Vázquez-Sánchez, "Non-destructive techniques based on eddy current testing," Sensors, vol. 11, no. 3, pp. 2525-2565, 2011.

[5] J. Prasongkit, A. Grigoriev, B. Pathak, R. Ahuja, and R. H. Scheicher, "Theoretical Study of Electronic Transport through DNA Nucleotides in a Double-Functionalized Graphene Nanogap," Journal of Physical Chemistry C, vol. 117, no. 29, pp. 15421-15428, 2013.

[6] C. Nú?Ez, P. A. Orellana, L. Rosales, R. A. R?Mer, and F. Domínguez-Adame, "Spin-polarized electric current in silicene nanoribbons induced by atomic adsorption," Phys.rev.b, vol. 96, no. 4, pp. 045403, 2017.

[7] "Noise in Graphene Superlattices Grown on Hexagonal Boron Nitride," ACS Nano, 2015.

[8] S. H. Kim, "Influence of sulfides on the tribological properties of composites produced by pulse electric current sintering," International Journal of Minerals Metallurgy and Materials, vol. 21, no. 1, pp. 95-103, 2014..

[9] F. A. de Souza, R. G. Amorim, W. L. Scopel, and R. H. Scheicher, "Controlled current confinement in interfaced 2D nanosensor for electrical identification of DNA," Physical Chemistry Chemical Physics, vol. 21, no. 45, pp. 24884-24890, 2019.

[10] A. Borsic, R. Halter, Y. Wan, A. Hartov, and K. D. Paulsen, "Electrical impedance tomography reconstruction for three-dimensional imaging of the prostate," Physiological Measurement, vol. 31, no. 8, pp. S1-S16, 2010.

[11] M. E. Spira, and A. Hai, "Multi-electrode array technologies for neuroscience and cardiology," Nature Nanotechnology, vol. 8, no. 2, pp. 83-94, 2013.

[12] S. Bhandari, G. H. Lee, A. Klales, K. Watanabe, T. Taniguchi, E. Heller, P. Kim, and R. M. Westervelt, "Imaging Cyclotron Orbits of Electrons in Graphene," Nano Letters, vol. 16, no. 3, pp. 1690, 2015.

[13] K. C. Nowack, E. M. Spanton, M. Baenninger, M. Koenig, J. R. Kirtley, B. Kalisky, C. Ames, P. Leubner, C. Bruene, and H. Buhmann, "Imaging currents in HgTe quantum wells in the quantum spin Hall regime," Nature Materials, vol. 12, no. 9, pp. 787-791, 2013.

[14] J.-P. Tetienne, N. Dontschuk, D. A. Broadway, A. Stacey, D. A. Simpson, and L. C. L. Hollenberg, "Quantum imaging of current flow in graphene," 2016.

[15] B. Du, Y. He, Y. He, and C. Zhang, "Progress and trends in fault diagnosis for renewable and sustainable energy system based on infrared thermography: A review," Infrared Physics & Technology, pp. 103383, 2020.

[16] J. Wu, J. Zhu, H. Xia, C. Liu, X. Huang, and G. Y. Tian, "DC-biased Magnetization Based Eddy Current Thermography for Subsurface Defect Detection," IEEE Transactions on Industrial Informatics, pp. 1-1, 2019.

[17] Y. He, B. Du, and S. Huang, "Noncontact Electromagnetic Induction Excited Infrared Thermography for Photovoltaic Cells and Modules Inspection," IEEE Transactions on Industrial Informatics, vol. 14, no. 12, pp. 5585-5593, 2018.

[18] Z. Zhao, X. Fan, G. Xu, L. Zhang, Y. Qi, and K. Zhang, "Aggregating Deep Convolutional Feature Maps for Insulator Detection in Infrared Images," IEEE Access, vol. 5, pp. 21831-21839, 2017.

CycleGANs for Semi-Supervised Defects Segmentation

Jiangshan Ai
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China, 611731
Jiangshan_ai@163.com

Sihua Chen
Shanghai Space Propulsion
Technology Research Institute
Shanghai, China
sihua2004@163.com

Peng Deng
Sichuan Aerospace Measurement
& Test Research Institute
Sichuan, China
xincunziran2004@163.com

Libing Bai *
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
bailb991@163.com

Lulu Tian
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
tl.110119@163.com

Jie Zhang
School of Automation Engineering
University of Electronic Science and
Technology of China
Chengdu, China
zhj06 _19@uestc.edu.cn

Abstract—**At present, many vision-based inspection methods are widely using for quality control in different fields. And the deep learning method has made a magnificent breakthrough in a variety of computer vision tasks, mainly through the use of largescale annotated datasets. Utilizing these progress is an option to improve defect segmentation performance. However, in the field of vision-based non-destructive testing (NDT), obtaining large scale annotated datasets is a great challenge. In this paper, a fully convolution neural network (FCN) is supervised trained using a small number of pixel-level annotated data for defect segmentation. Simultaneously, Cycle-Consistent Generative Adversarial Networks (CycleGANs) are used to learn the segmentation in an unsupervised way as a supplement. The requirement of annotated data is then reducing by utilizing many un-annotated data. Experiments on the published GDXray dataset show that the framework based on CycleGANs is effectiveness for defect image segmentation using only a few labelled samples.**

Index Terms—CycleGANs, Image Segmentation, Nondestructive Testing, Defect Inspection, Semi-supervised Learning

I. INTRODUCTION

With the fast development of social industrialization, abundant of inspection methods have been introduced for better quality control. Testing has undoubtedly been recognized as a very important link in the process of product realization. Accurate detection by Non-destructive testing is proved and preferred to be adopted in various industries due to its advantages of lean and efficient. Meanwhile, neural network is widely used for feature extraction and image segmentation, therefore, they are very suitable for vision-based NDT defects segmentation. And the continuous development of deep learning makes the convolutional neural networks (CNNs) the most common tools used for image segmentation. Uses large scale labelled image datasets for training, the CNN-based image segmentation methods have achieved outstanding results in various image segmentation tasks [1] [2]. However, it is still a great challenge to train deep neural networks with a limited number of annotated data. In the vision-based NDT field, a common problem is lack of large-scale annotation datasets. Training a deep neural network with only a few labelled data will result in the overfitting problem. To overcome overfitting, a regularization term is typically introduced to improve the deep networks' generalization capabilities.

In this paper, we aim to find out a solid way to overcome the problem by introducing unsupervised learning based on CycleGANs [3] as a regularization term. In this situation, we need only a few labelled defect images to do supervised learning and use many unlabelled defect images to do unsupervised learning simultaneously. In this setting, we have two subsets of the training data: $S_s = \{(x_i, y_i)\}_{i=1}^n$ contains labelled defect images x_i and their corresponding segmentation labels y_i, and $S_u = \{x_j\}_{j=1}^m$ contains unlabelled defect images. To date as of today, CycleGANs is more popular to be used for image-to-image translation (also related to domain adaption) in comparison with other traditional translation methods, because the unpaired images can be also adopted and applied rather than paired images only. Since our purpose is not to have a very high-performance segmentation network but to solve the limitation of the number of annotated data problem, we just adopt a commonly used fully convolutional neural network as the segmentation networks, which is also part of the CycleGANs. We use the subset S_s to supervised train the translation from defect images to corresponding segmentation labels, and simultaneously unsupervised train the bidirectional mapping from the subset S_u to $\{y_i\}_{i=1}^n$ in subset S_s.

We build our training and testing datasets by extracting ROI images with the defects and their corresponding segmentation labels from the welding part of the GDXray dataset [4]. We use a small proportion of training dataset for supervised training, and the left training defect images are used for unsupervised training. To get a performance standard to compare with, we use all the training data to train the segmentation network in a fully supervised way.

The contributions of this paper are listed as follows:

We realize to well train a deep FCN to do defect segmentation with a few pixel-level labelled defect images by applying the CycleGANs with many unlabelled images.

---

*Libing Bai is the corresponding author. (E-mail: bailb991@163.com).

With the application of our proposed method, defects segmentation performance by using a few labelled images is comparable with the performance by using a big number of labelled images.

The overall idea of the paper is to use CycleGANs-based domain adaption to learn the mapping from defect images to defect segmentation by utilizing the unlabelled defect images which will reduce the annotation workload, and make training a deep defect segmentation neural networks possible with only limited labelled defect images. In the following, we briefly describe the relevant research field of image segmentation and GANs in section 2. Then the methodology and experiments are presented in section 3 and 4 respectively. Finally, the conclusion and future works are shown in section 5.

II. RELATED WORK

A. Image Segmentation

Performance of image semantic and instance segmentation has achieved significant progress via CNNs based supervised learning methods [5] [6]. Considering instance segmentation, there are mainly two pipelines, one is detecting object ROIs first and then segment them [7], the other is to do semantic segmentation first and then use post-process to identify different objects in each class [8]. For semantic segmentation, only the classes' segmentations are needed [1], here we only consider this scenario. In spite of this progress, a big number of pixel-level labelled images are still required to train the networks, which are very costly to acquire. One way to overcome the restriction is to use weakly-supervised methods via labelled images which are easier to achieve, like imagelevel annotation [9], bounding boxes [10], etc. However, a certain amount of labelled images are still acquired for weakly supervised training. In this article, a semi-supervised approach is adopted to minimize the training data acquirement cost. Semi-supervised learning is intended to use many unlabelled images via unsupervised training as supplementation for the supervised training of a few labelled images. In recent time, variant of semantic segmentation methods via semi-supervised learning approach have been proposed, for example self-learning [11], manifold embedding [12], distillation [13] etc. There are a few CycleGANs based semi-supervised image segmentation researches in other fields, mainly for daily life images and medical images [14].

B. Generative Adversarial Networks

Generative adversarial nets (GANs) [15] is a special differentiable generative model which is composed of a pair of networks and can be trained using unlabeled data. It is like a two-player game, one is G (z), and the other is D(x). G (z) is a generator, D(x) is a discriminator. G (z) is trying to generate fake data (for instance images) from a random noise z, with the aim of making realistic images. D(x) receives both the real images and the generated images as x and aims to differentiate them. Both G (z) and D(x) are trained simultaneously, and in competition with each other. The main motivation behind GANs is to capture the density representation of data from the training dataset implicitly.

Recently, the GANs have got great attention in the computer vision field for their ability to generate high-quality synthetic images and they can be trained in an unsupervised way. GANs have been successful in lots of image generation and domain adaptive applications, such as super resolution [16], text-to-image synthesis [17], image in painting [18], image-to-image translation [19], and domain adaption [20]. GANs based domain adaption [21] usually uses adversarial training to match the distribution of source and target in data or feature space. Normally, paired-samples (samples of same content for source and target domains) which are more expensive than unpaired samples (samples of different content for source and target domains) to acquire are needed for training. The two variants of the GANs, Cycle-Consistent Adversarial Networks (CycleGANs) and Coupled GANs [22] are proposed so that the translation can be done using unpaired images which is more efficient for doing domain adaption [23].

III. METHODOLOGY

A. CycleGANs for Semi-supervised Defect Segmentation

The proposed defect segmentation model is based on the CycleGANs, which is popular for image-to-image translation using unpaired data. In the original CycleGANs, there are two pairs of generator and discriminator, and each pair is trained in the adversarial learning way. One generator is used to learn a translation from image domain X to the other image domain Y, and the corresponding discriminator is employed to identify the images which are real from Y or fake generated from X. The other generator is used to learn a backward translation from Y to X, and the paired discriminator is employed to identify the images which are real from X or fake generated from Y. The generators learn bidirectional mapping by spoofing the discriminators. For using the unpaired training images, the bidirectional cycle consistency losses are adopted so that the images from X mapped to Y and then mapped back to X (recovered) have less difference and the same for the images from Y mapped to X back to Y.

As in the original CycleGANs, our model also has four parts, for combining the supervised and unsupervised learning, we use one generator which generates segmentation label images from the defect images as the segmentation network. The proposed framework for defect images segmentation is shown in Fig.1, one of the two generators G_{LD} maps the labels to the defects, and the other one G_{DL} maps the defects backwards to labels. The two adversarial discriminators D_D discriminate between real defects and fake defects, and D_L discriminate between real labels and fake labels. And the three types of losses will be discussed in the following.

B. Loss Functions

There are three distributions: the labelled defect images $x \sim p_{data}(x)$, the corresponding segmentation labels $y \sim p_{data}(y)$ and the unlabelled defect images $x' \sim p_{data}(x')$. Our goal is to learn the mapping from domain X to Y, the training samples $\{x_i\}_{i=1}^N$, $\{x'_j\}_{j=1}^M$ where $x_i \in X, x'_j \in X$ and $\{y_i\}_{i=1}^N$ where $y_i \in Y$. As shown in Fig.1, there are several losses to our model. The adversarial (GAN) losses are
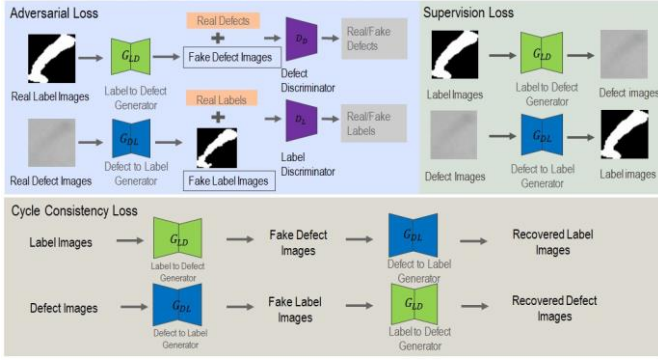
Fig. 1. The schematic framework of the defects segmentation networks

utilized for training the generators and discriminators in an alternative way. The original cross-entropy loss was replaced by the least-squares loss [24] which gives more stable training. The first term as in formula (1) is used to update the G_{LD} so that it can generate more realistic defects. The second term in formula (2) is used to update the G_{DL} for generating more realistic segmentation label images. The formula (3) and the formula (4) are the discriminator losses for updating the D_D and D_L respectively.

$$L^D_{G_{LD}}(G_{LD}, D_D) = \mathbb{E}_{y \sim p_{data}(y)}[(D_D(G_{LD}(y)) - 1)^2] \quad (1)$$

$$L^L_{G_{DL}}(G_{DL}, D_L) = \mathbb{E}_{x' \sim p'_{data}(x')}[(D_L(G_{DL}(x')) - 1)^2] \quad (2)$$

$$L^D_{D_D}(G_{LD}, D_D) = \mathbb{E}_{y \sim p_{data}(y)}[(D_D(G_{LD}(y)))^2] + \\ \mathbb{E}_{x' \sim p'_{data}(x')}[(D_D(x') - 1)^2] \quad (3)$$

$$L^L_{D_L}(G_{DL}, D_L) = \mathbb{E}_{x' \sim p'_{data}(x')}[(D_L(G_{DL}(x')))^2] + \\ \mathbb{E}_{y \sim p_{data}(y)}[(D_L(y) - 1)^2] \quad (4)$$

Combining formula (1) and (3) as the adversarial losses for defect images, and the same for formula (2) and (4) as the adversarial losses for label images. We have one GAN loss objective:

$$L_{GAN}(G_{LD}, G_{DL}, D_D, D_L) = L^D_{GAN}(G_{LD}, D_D) + \\ L^L_{GAN}(G_{DL}, D_L) \quad (5)$$

The Cycle-Consistency Loss is added for using unpaired training data which is suitable for our using case. The L1 norm is used to calculate the reconstruction losses for both labels and defects for sharper generated images. Formula (6) is the cycle loss from labels to defects to labels, and formula (7) is the cycle loss from defects to labels to defects. Written together we have a total Cycle-Consistency loss formula (8).

$$L^{LDL}_{cycle}(G_{LD}, G_{DL}) = \mathbb{E}_{y \sim p_{data}(y)}[||(G_{DL}(G_{LD}(y)) - y||_1] \quad (6)$$

$$L^{DLD}_{cycle}(G_{DL}, G_{LD}) = \mathbb{E}_{x' \sim p'_{data}(x')}[||G_{LD}(G_{DL}(x')) - x'||_1] \quad (7)$$

$$L_{cycle}(G_{LD}, G_{DL}) = L^{LDL}_{cycle}(G_{LD}, G_{DL}) + \\ L^{DLD}_{cycle}(G_{DL}, G_{LD}) \quad (8)$$

The above losses are used for unsupervised learning. For supervised learning, we use the annotated ground truth labels as supervision to the generator G_{DL} which is also our segmentation network, and the loss function is formula (9). Since for generator G_{LD}, the labelled defect images is the generating target, we add another supervision loss in formula (10). The combined supervision loss is in formula (11).

$$L^L_{sup}(G_{DL}) = \mathbb{E}_{x,y \sim p_{data}(x)p_{data}(y)}[||G_{DL}(x) - y||_1] \quad (9)$$

$$L^D_{sup}(G_{LD}) = \mathbb{E}_{x,y \sim p_{data}(x)p_{data}(y)}[||G_{LD}(y) - x||_1] \quad (10)$$

$$L_{sup}(G_{LD}, G_{DL}) = L^L_{sup}(G_{DL}) + L^D_{sup}(G_{LD}) \quad (11)$$

Finally, combining the GAN loss (5), the cycle loss (8) and the supervision loss (11) with two adjustable super-parameters λ_1 and λ_2, we have the full objective:

$$L_{total}(G_{LD}, G_{DL}, D_L, D_D) = L_{GAN}(G_{LD}, G_{DL}, D_L, D_D) + \\ \lambda_1 L_{cycle}(G_{LD}, G_{DL}) + \\ \lambda_2 L_{sup}(G_{LD}, G_{DL})$$

(12)

Our optimizing target is to find the optimal G^*_{LD} and G^*_{DL} by solving the following formula:

$$arg \min_{G_{LD}, G_{DL}} \max_{D_L, D_D} L_{total}(G_{LD}, G_{DL}, D_L, D_D) \quad (13)$$

IV. EXPERIMENTS AND RESULTS

A. Dataset

There are very few public large-scale datasets in the vision based NDT domain, and it is difficult to build qualified largescale datasets. In our paper, the GDXray dataset is used, which consists of 19,407 X-ray images from different aspects.

What we use is the welding part which consists of 88 images

with 641 labelled bounding boxes of the welding defects ROIs and corresponding segmentation labels of these defect images. Fig.2 shows some examples of datasets we are using in this paper.
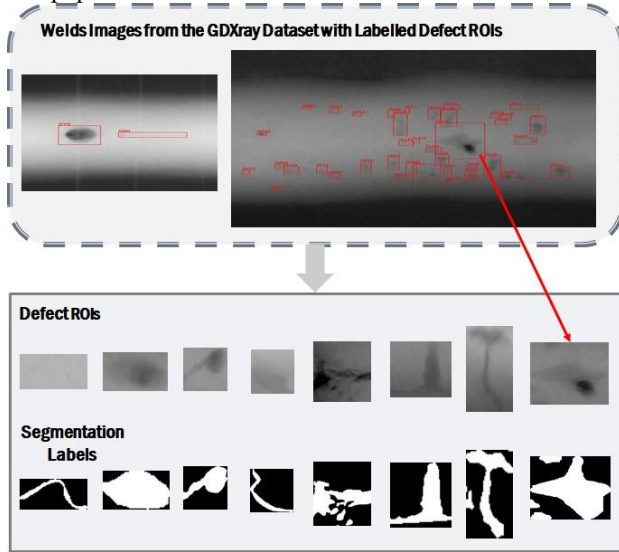

Fig. 2.The examples of the welds part of the GDXray dataset.

We use 640 ROI images in the welding dataset (drop the too-small one) as the basis, then randomly crop the images into many pieces by making sure there is enough proportion of defect and background in the cropped images, and we have 2000 defect images. 1600 images are utilized in the training dataset, and the left 400 images are used as the testing dataset. Then we resize all the images to 128 × 128. For all the images, we make use of traditional data augmentation pipeline of random cropping and flipping to augment the dataset before training.

B. Evaluation Standard

Here the mean intersection over union (mIoU) is used as an evaluation metric of the segmentation results. It is a mean value of IoU for all the classes in the images. The IoU metric for defect segmentation is defined as $TP/(TP+FP+FN)$ where the TP, FP, FN and related TN are interpreted and presented as follows:

True Positive (TP): number of defect pixels correctly classified.

True Positive (TN): number of no-defect pixels correctly classified.

False Positive (FP): number of no-defect pixels classified as defects pixels.

False Positive (FN): number of defect pixels classified as no-defect pixels.

C. Implementation Details

Since the generators in the original CycleGANs are already fully convolution networks which are able to do the image segmentation, we only use the original network architecture, although other segmentation networks can be used, and should be able to improve the segmentation results. The original generator architecture which follows generative

networks proposed in Johnson et al. [25] has got remarkable results in image-to-image translation tasks. This network
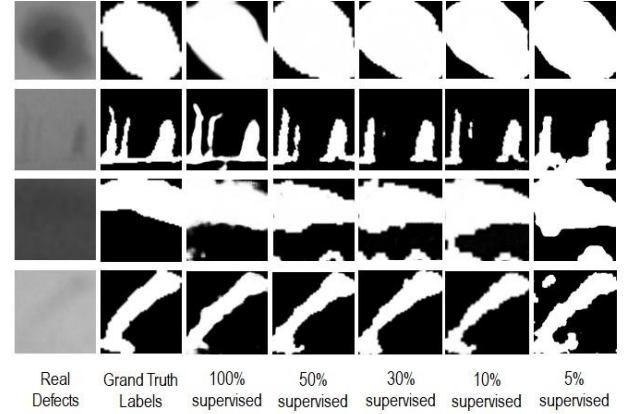
Fig. 3. Examples of defect images, ground truth labels and segmentation results using different percentage of supervised training.

has one input convolution block, two down sampling convolution blocks, nine residual blocks, two up sampling convolutions blocks, and one output convolution block. Similarly, instance normalization [26] is used in all the convolution blocks except for the output block. Furthermore, the tanh is adopted as the output function for the two generators. In preprocessing, each grayscale image is randomly horizontally flipped and normalized to [-1, 1].

Follow the original CycleGANs model, the 70 × 70 patch discriminators [19] are used. This is realized by design suitable convolutional blocks so that the output receptive field for each output pixel is 70. The patch discriminators perform better than pixel-level discriminators or image-level discriminators. The discriminator is composed of five convolutional blocks, also the instance normalization is used, and all the convolution blocks are using the Leaky ReLU activations except for the last convolutional layer which is followed by average pooling. Two Adam optimizers [27] with $\beta1 = 0.5$ and $\beta2 = 0.999$ are used for optimizing the networks. The learning rate is initially set as 2×10^{-4}, and can liner decay after several epochs during the training. The batch size is 6 in all training and 1 in all testing. The values of the super-parameters in formula.(12) are set to $\lambda1 = 10$, and $\lambda2 = 5$.

D. Evaluation of Results

Fig.3 displays some examples of the segmentation results from the testing dataset by using different percentage of labelled defects images from the training dataset for supervised training, and all the data in the training dataset are used for unsupervised learning. From the visual perception, even using a small proposition data to supervised train the segmentation network, we can get comparable segmentation images from the unsupervised learning. The quantitative segmentation performances of mIoU are shown in Table 1. With the increasing number of data for supervised learning, the performance improves slightly, and we get the up-bound performance by using all the training data for supervised learning. From the above results, in supervised learning, the more labelled data used the better performance we get. But in the case of only a few labelled training data are available, the proposed method can be applied to improve the segmentation

results in the situation that many unlabelled data can be acquired.

TABLE I
DIFFERENT PERCENTAGE OF SUPERVISION

Labelled (%)	mIoU
100	0.7118
50	0.6646
30	0.6600
15	0.6532
5	0.6358

The above experiments show the influence of the changing number of supervised learning data with the same number of all available defect images as unsupervised learning data. To further present the effectiveness of using the unlabelled defect images to improve the segmentation network's performance, we also use the fixed number of labelled data during training about 5% for supervised learning with the different number of unlabelled data used for unsupervised learning. Table 2 demonstrates the results of these experiments. Using only supervised learning with the 80 defect images (5% from the training dataset) and their corresponding labels, we get the worst segmentation results. By using the proposed method with the increasing number of unlabelled defect images, the segmentation performance is improved gradually. From the mIoU number, the improvement is not very big, this may be because the defect dataset is simple and using a small number of training data with traditional augmentation can have a good result to some extent.

We also test using the cross-entropy loss for both the supervised segmentation label loss and the recovered label loss, which gives similar results as using the L1 norm losses. Since both the label images and defect images usually have large gradients along the defect contour, we test several similarity measures of the image gradients for further guarantee the consistency of image content in the two domains. However, the segmentation performance only improves slightly, that we will further optimize later.

TABLE II
DIFFERENT PERCENTAGE OF UNSUPERVISED LEARNING

Unlabelled(%)	mIoU
100	0.6358
80	0.6254
50	0.6232
0	0.6158

V. Conclusion and future works

In this paper, a semi-supervised defect image segmentation method based on CycleGANs is proposed. This method uses a large number of unlabeled defect images to reduce the requirement for labelled defect images. We apply this technique
on the welding defects of GDXray dataset, and the results show that the segmentation performance of the network is improved with the increasing number of unlabeled images used for training, which means the effectiveness of the method.

Now only the welding defect images of the GDXray dataset are tested, later we will test other datasets, and try some different method to improve the segmentation performance, such as using other types of segmentation network architectures, adding extended content metric loss to regulate the unsupervised learning or utilizing the attention mechanism to focus the learning on the defects. Moreover, from the experiments, except for our segmentation generator, the other generator is learning to generate fake defect images which might be used for synthetic defects data augmentation. And we expect that using the proposed semi-supervised method is not only beneficial to the semantic segmentation tasks but also to the instance segmentation tasks which we will test in the future.

References

[1] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," pp. 3431–3440, 2015.

[2] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[3] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," pp. 2242–2251, 2017.

[4] D. Mery, V. Riffo, U. Zscherpel, G. Mondragon, I. Lillo, I. Zuccar, H. Lobel, and M. Carrasco, "Gdxray: The database of x-ray images for nondestructive testing," *Journal of Nondestructive Evaluation*, vol. 34, no. 4, p. 42, 2015.

[5] L. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoderdecoder with atrous separable convolution for semantic image segmentation," pp. 833–851, 2018.

[6] H. Chen, K. Sun, Z. Tian, C. Shen, Y. Huang, and Y. Yan, "Blendmask: Top-down meets bottom-up for instance segmentation," *arXiv: Computer Vision and Pattern Recognition*, 2020.

[7] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask r-cnn," pp. 2980– 2988, 2017.

[8] A. Arnab and P. H. S. Torr, "Bottom-up instance segmentation using deep higher-order crfs," *arXiv: Computer Vision and Pattern Recognition*, 2016.

[9] G. Papandreou, L. Chen, K. Murphy, and A. L. Yuille, "Weakly- and semi-supervised learning of a dcnn for semantic image segmentation," *arXiv: Computer Vision and Pattern Recognition*, 2015.

[10] J. Dai, K. He, and J. Sun, "Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation," pp. 1635– 1643, 2015.

[11] W. Bai, O. Oktay, M. Sinclair, H. Suzuki, M. Rajchl, G. Tarroni, B. Glocker, A. P. King, P. M. Matthews, and D. Rueckert, "Semisupervised learning for network-based cardiac mr image segmentation," pp. 253–260, 2017.

[12] C. Baur, S. Albarqouni, and N. Navab, "Semi-supervised deep learning for fully convolutional networks," pp. 311–319, 2017.

[13] Y. Zhou, Y. Wang, P. Tang, S. Bai, W. Shen, E. K. Fishman, and A. L. Yuille, "Semi-supervised 3d abdominal multi-organ segmentation via deep multi-planar co-training," pp. 121–140, 2019.

[14] A. K. Mondal, A. Agarwal, J. Dolz, and C. Desrosiers, "Revisiting cyclegan for semi-supervised segmentation." *arXiv: Computer Vision and Pattern Recognition*, 2019.

[15] I. J. Goodfellow, J. Pougetabadie, M. Mirza, B. Xu, D. Wardefarley, S. Ozair, A. C. Courville, and Y. Bengio, "Generative adversarial nets," pp. 2672–2680, 2014.

[16] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. P. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single

image super-resolution using a generative adversarial network," *arXiv: Computer Vision and Pattern Recognition*, 2016.

[17] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," *arXiv: Neural and Evolutionary Computing*, 2016.

[18] R. A. Yeh, C. Chen, T. Y. Lim, A. G. Schwing, M. Hasegawajohnson, and M. N. Do, "Semantic image inpainting with deep generative models," pp. 6882–6890, 2017.

[19] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," pp. 5967–5976, 2016.

[20] Y. Tsai, W. Hung, S. Schulter, K. Sohn, M. Yang, and M. Chandraker, "Learning to adapt structured output space for semantic segmentation," pp. 7472–7481, 2018.

[21] J. Hoffman, D. Wang, F. Yu, and T. Darrell, "Fcns in the wild: Pixel-level adversarial and constraint-based adaptation." *arXiv: Computer Vision and Pattern Recognition*, 2016.

[22] M. Y. Liu, T. Breuel, and J. Kautz, "Unsupervised image-to-image translation networks," 2017.

[23] J. Hoffman, E. Tzeng, T. Park, J. Zhu, P. Isola, K. Saenko, A. A. Efros, and T. Darrell, "Cycada: Cycle-consistent adversarial domain adaptation," *arXiv: Computer Vision and Pattern Recognition*, 2017.

[24] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," 2016.

[25] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II*, ser. Lecture Notes in Computer Science, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., vol. 9906. Springer, 2016, pp. 694–711. [Online]. Available: https://doi.org/10.1007/978-3-319-46475-6 43

[26] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," *arXiv: Computer Vision and Pattern Recognition*, 2016.

[27] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv: Learning*, 2014.

BP-Neural-Network-Based Aging Degree Estimation of Power Transformer Using Acoustic Signal

Yukun Zhang
School of Physic & Information Technology
Shaanxi Normal University
Xi'an, China
zy_kun@snnu.edu.cn

Jisheng Li[*]
School of Physic & Information Technology
Shaanxi Normal University
Xi'an, China
lj_sheng@snnu.edu.cn

Hanchao Liu
School of Physic & Information Technology
Shaanxi Normal University
Xi'an, China
lhc1159391127@163.com

Rong Liu
School of Physic & Information Technology
Shaanxi Normal University
Xi'an, China
liu_rong_007@snnu.edu.cn

Fan Yang
School of Physic & Information Technology
Shaanxi Normal University
Xi'an, China
yangfan1997a@163.com

Ting Li
School of Physic & Information Technology
Shaanxi Normal University
Xi'an, China
lt15635593689@163.com

Abstract—**In this paper, an aging degree estimation method using acoustic signals is proposed based on a BP neural network. Twenty-eight transformers are taken as research objects. The transformer's internal noise mechanism is analyzed, and the acoustic signals of the high- and low-voltage sidewalls are collected and screened. The BP neural network is used to predict the transformer age in real-time. Comparing the predicted results with their actual operation time provides a sufficient basis for determining the degree of transformer aging and the need for an overhaul. After network training and data testing, the error between the predicted value and the actual value reaches the least. The proposed estimation method can play an innovative role in the process of transformer fault monitoring.**

Keywords—acoustic signal; BP neural network; transformer age; fault monitoring

I. INTRODUCTION

Transformers are one of the most critical pieces of equipment in power systems. These electrical equipments gradually develop defects due to mechanical aging and heat aging during long-term operation, which in severe cases, can affect the stable operation of the entire power system [1]. Many transformers experience severe failures in distribution networks that do not meet power requirements before they reach the design life [2–4]. According to transformer statistics, over time, the primary sources of accidents are associated with the iron cores and windings. Even when a transformer continues to operate, it may contain certain intrinsic defects that degrade the short-circuit resistance.

Researchers have focused on the transformers' aging degree and their critical state in real-time to prevent the conversion of the fault state in advance effectively and to reduce or avoid losses caused by the internal structure deterioration of power transformers [5,6]. After the long-term operation, the transformer's commissioning time can indirectly indicate the degree of internal aging. However, the degrees of aging of the transformer under different working conditions are different, so it is necessary to perform an accurate evaluation based on the transformer's specification and operating conditions. In the process of long-term online monitoring of the transformer's operation status, the corresponding electrical characteristics also change with internal aging [7–9].

The transformer's noise is closely related to its mechanical properties and is an essential technical parameter for the state evaluation. The magnitude of the transformer's noise level is one of the essential indicators to measure its operation level and defect degree. The noise is generated mainly by the core and the winding vibration. The core vibration comes mainly from the magnetostrictive force and the magnetic leakage's electromagnetic force caused by the silicon steel sheet in the magnetic field. The winding generates a leakage flux after the current flows through, and the winding vibrates due to exposure to magnetic field leakage. In the long-term vibration, the silicon steel sheet inevitably loosens, or the iron core is slightly deformed, and the magnetostriction is intensified; the winding thus becomes loose, deformed, or even damaged, and the vibration increases [10-12]. Whether the transformer's internal structure is intact, mixed noise is transmitted to the outer casing through the structural components and radiated to the surroundings, so the surface noise of the transformer can be used as an essential factor in evaluating the condition of the device.

Many studies have been carried out worldwide by analyzing power transformers' acoustic characteristics through acoustic measurements [13-15]. The vibro-acoustic signals of power transformers have been measured in the field, and the time-frequency domain characteristics and noise levels have been analyzed. In [16], the noise level and spectrum characteristics of 1000 kV transformers in Shaanxi Province of China were measured and analyzed in detail. The above experimental research shows that it is feasible to monitor transformers online based on vibro-acoustic signals [17-19]. Combined with the above research status, the following research work has been done in this paper. The collected transformer acoustic signals were processed, and the operation time was predicted by the BP neural network and compared with its actual investment period. If the deviation between the predicted operating time

*Jisheng Li is the corresponding author. (e-mail: lj_sheng@snnu.edu.cn)

and the actual operation time is large, it indicates that the transformer's aging speed is greater than the natural trend. The accelerated aging has a continuous impact on the transformer's internal structure. Therefore, the method has innovative significance for the timely detection of internal and other hidden defects in transformers.

II. Noise Transmission Path And Acquisition Mode

A. Noise Transmission Path

The transformer's noise emitted through the air is composed of two parts. One part is the cooling device noise caused by the cooling fan's vibration and the transformer's oil pump. The other part is intrinsic noise generated by the primary vibration sources, such as the core and winding, and transmitted to the tank wall (including the magnetic shield). One of the transmission paths of this noise component is through the foundation of the core bracket to the tank: the solid path; the second path is through the insulating oil to the tank, that is, along the fluid path.

The magnetostrictive effect of silicon steel sheets and the electromagnetic force between silicon steel sheets are the main reasons in the intrinsic noise. On the one hand, electromagnetic attraction occurs between the joint of the silicon steel sheet and the laminated sheet due to magnetic leakage. This effect also causes some noise, but this noise component can be neglected because of the improved technology of making silicon steel sheets. On the other hand, when the core is magnetized, the silicon steel sheet's size increases along the direction of the magnetic force line, and the size of the silicon steel sheet perpendicular to the direction of the magnetic force line decreases.

B. Acquisiton Mode

The transformer acoustic acquisition system, shown in Fig. 1, includes microphones, acquisition device, and a terminal computer. The acoustic acquisition system can collect, analyze and save multi-channel acoustic signals synchronously, and carry out real-time ground spectrum analysis of the current signal, and set threshold alarm if necessary.
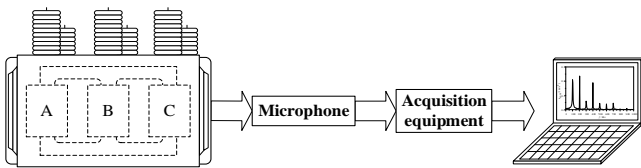
Fig. 1. Acoustic signal acquisition system

TABLE I. Acoustic Sensor Parameters

Sensor Parameters	Acoustic Sensor
Sensitivity	40mV/Pa
Range	16~134dB
Installation Mode	Bracket Support
Frequency Band	20~20kHz
Application Temperature	-30℃ ~ 80℃

Table 1 shows the microphone parameters. The microphone is supported by a wheat rack and placed near the transformer tank. In this paper, multiple microphones

are used to collect the transformer's acoustic signals synchronously, and then the collected signals are transmitted to the terminal computer.

III. Aging Degree Prediction Model

The back propagation neural network is one of the most successful neural network learning algorithms. Influenced by multiple nonlinear external factors, the network can learn and store a large number of input-output mode mapping relationships without revealing the mathematical equation describing the mapping relationship beforehand. In recent years, BP neural networks have been used in transformer fault diagnosis. The network model topology includes the input layer, hidden layer, and output layer, as shown in Fig. 2.

Each neuron in the upper layer is connected to each neuron in the lower layer, but there is no connection between the left and right neurons. When a pair of learning modes are provided to the network, the neurons' activation values in each layer propagate from the input layer through the hidden layers to the output layer. Then, according to the principle of reducing the error between the ideal output and the actual output, the network corrects each connection weight layer by layer from the output layer to the input layer. During the iteration of the network, as each layer's weight values are updated, the correct rate of the network response to the input mode will also be continuously improved.

In this paper, the function recurrence capability of the neural network is used to predict the year of operation, taking the acoustic signal on the transformer tank's surface as the input vector and the operation time as the output vector.

A. Sample Collection and Selection

The original data of transformers come from some substations in China. In order to improve the accuracy of signal acquisition, the microphone is symmetrically arranged. The large transformer windings are placed on the base, although part of the vibration energy is transmitted to the tank wall through cooling oil, most of the vibration energy is transmitted mainly from the bottom to the top through the structural parts. Due to the attenuation in vibration energy in the transmission process, the measuring point should be as close as possible to the transformer's bottom.

Eight microphones with stands are placed symmetrically on the high-voltage and low-voltage side, so the sampling
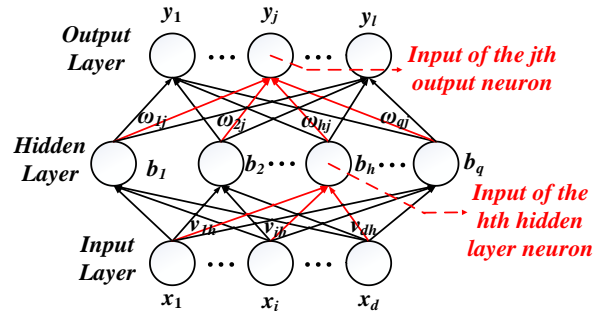

Fig. 2. Topological structure of network model

Fig. 3. Data acquisition site

data of the eight channels are obtained, and the sampling frequency of acoustic data is 10240 Hz. As shown in Fig. 3, the horizontal distance between each microphone and the oil tank surface is 0.1 m, and the height is approximately 1/3 the height of the tank.

The time-domain waveforms of acoustic signals at various points are different. Fig. 4 shows the acoustic time-domain signals collected at four different locations. Each waveform has its own characteristics, so many measuring points can collect as many kinds of time-domain signals as possible. It is with multi-directional noise acquisition that
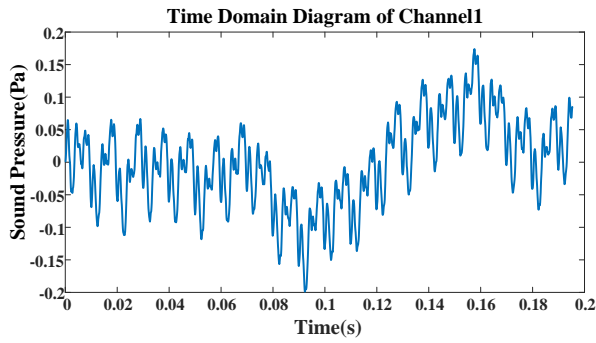
(a) Channel 1
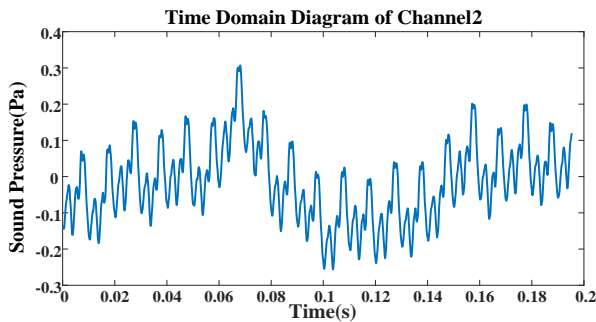
(b) Channel 2
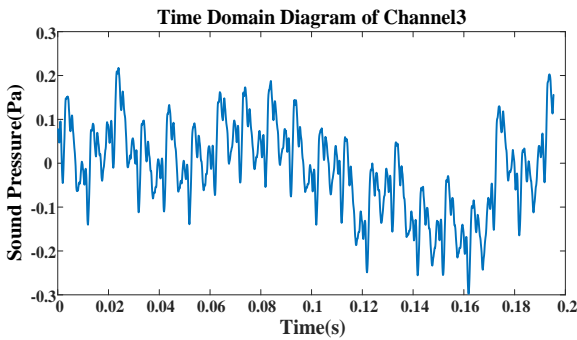
(c) Channel 3
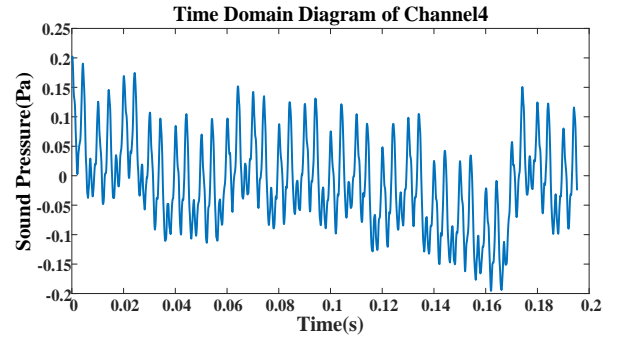
(d) Channel 4

Fig. 4. Acoustic signal

transformers with different aging levels can be evaluated more accurately.

The appropriate acoustic data capacity should be chosen before training the model. If the data capacity is too large, the correlation between individuals is strong, the whole model's stability will be affected. The oscillation of the neural network is not conducive to the convergence of the function. If the amount of data is too inadequate, it is not easy for the neural network to achieve the fitting effect and the prediction accuracy. In this paper, the acoustic signals of transformers with the same voltage level are collected and filtered, and representative data are selected to train the network.

B. Data Preprocess

The input data are normalized and transformed into [0,1]. The transformation formulas are as follows:

$$y = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

The normalization of samples is closely related to the convergence rate of learning. The purpose of scaling transformation for different feature dimensions is to make the influence weights consistent for each feature dimension on the objective function. Experience has proven that data normalization can significantly improve the speed and accuracy of learning. The output of the network is inversely normalized, and the formula is as follows:

$$x = y * (x_{max} - x_{min}) + x_{min} \tag{2}$$

In the process of training network, the network is trying to minimize the error. The normalization makes the optimization problem can reach the destination quickly and accurately.

C. Network Model Structure

In this prediction model, the number of neurons in the input layer is 103, the number of output cells is 1, the number of hidden layers is 3, and the numbers of neurons in the three hidden layers are 30, 20, and 10. From the substations, 30 sets of transformer acoustic data were collected. After the unnatural data were removed, 28 records were obtained, including a transformer with apparent defects. 24 transformers in good condition were selected as the training set, and the others were used as the testing data.

Each neuron's weights are initialized before the first iteration, and then the neural network is trained using a training set containing 24 transformer acoustic signals and

operation time. Each iteration of the network model makes the following calculations:

$$\begin{cases} net_i^{(l)} = \sum_{j=1}^{s} W_{ij}^{(l)} h_j^{(l-1)} + b_i^{(l)} & (3) \\ h_i^{(l)} = f\left(net_i^{(l)}\right) & (4) \end{cases}$$

where $h_j^{(l-1)}$ is the input sources of l-1 layer neurons to l layer neurons, and $h_i^{(l)}$ is the output of l layer neurons.

Before the network reaches the preset number of iterations, the loss function value obtained by the current output results and the actual values will be compared with the error threshold. When the loss function value is not less than the error threshold, the network will continue to revise the weight value in the next iteration; when the loss function value is less than the error threshold, the network will stop training. If the loss function value is still not lower than the error threshold after the preset number of iterations, the network will stop training. The loss function of the network in this model is expressed as follows:

$$E = \frac{1}{2m} \sum_{i=1}^{m} \left(d(i) - y(i)\right)^2 \qquad (5)$$

where $d(i)$ is the i-th desired output of the network, $y(i)$ is the i-th actual output of the network, and m is the number of outputs. After the network training is completed, another transformers' acoustic signals are input into the network to obtain the predicted results.

IV. RESULT ANALYSIS AND MODEL EVALUATION

A. Result Analysis

In the learning process, the neural network's error curve tends to be flat at the sixth iteration, and the error precision is maximized. The best mean square error is 3.45e-3.

Table 2 shows the predicted results and actual values of testing samples. Three transformers put into operation in 2006, 2007, and 2010 achieved good prediction results through the neural network. Considering the influence of temperature, humidity, and other factors, the error between the predicted value and the actual value should be less than four years. The transformer operated in 2012 is known to have defects inside, the predicted value and the actual value are at least seven years apart; that is, the transformer is actually in a more severe state than the state under the natural trend. Such a transformer should attract attention and be addressed by combining other methods for fault diagnosis.

TABLE II. RESULTS COMPARISON

Operation Time	2012	2010	2006	2007
Prediction	2004.2	2008.2	2010.0	2009.7
Error	-7.8	-1.8	2.9	2.7

The estimation accuracy of the model is high, allowing the goal to be reached of predicting the commissioning period of the transformer, with practical significance for evaluating the aging degree of the transformer. However, The training set is limited to testing the transformer of the same voltage level. The aging degree of transformers with other voltage levels needs further study.

B. Model Evaluation

The model is evaluated from three aspects: coefficient of determination (R-square), root mean square error (RMSE), and relative error (RE). The evaluation results are shown in Table 3. The coefficient of determination characterizes the fit of a model by the change in the data. The closer the coefficient is to 1, the stronger the explanatory power of the equation's input is, and the better the model fits the data. According to the usual experience, a value greater than 0.4 indicates an excellent fitting effect. The RMSE can reflect the measurement's precision well and is used to measure the deviation between the observed value and the actual value. The smaller the RMSE is, the higher the measurement accuracy. In general, the RE is more reflective of the prediction's credibility degree than the RMSE. The same unit dimension reflects the value of the prediction result deviating from the transformer's actual age. This parameter precisely indicates the true magnitude of the deviation from the actual value.

TABLE III. MODEL EVALUATION RESULTS

Model Evaluation Standard	R-square	RMSE	RE
Evaluation Result	0.56	2.99	8%

V. CONCLUSIONS

(1) The equipment parameters of the acoustic acquisition system used in this paper can meet the field requirements, and the fluctuation trend of acoustic signal at different measuring points is not the same.

(2) A BP neural network is used to predict the operation year of the transformer. After training the network and testing, the prediction model has high prediction accuracy for the transformer's operation time.

(3) Considering the impact of the onsite environment (e.g., temperature and humidity) on electrical equipment, the actual years and predicted years allow for three years of error.

(4) By comparing the model output with the actual value, the aging degree estimation method proposed in this paper can effectively distinguish one transformer with internal defects from three regular transformers, which provides a new idea for fault diagnosis of power equipment.

(5) The research in this paper also proves that deep learning has strong practicability in dealing with power equipment signals to some extent, and has a significant advantage in power equipment status evaluation and fault diagnosis.

REFERENCES

[1] Hui, Song, et al. "Power Transformer Operating State Prediction Method Based on an LSTM Network." *Energies* 11.4(2018):914.

[2] Soltanbayev, Ilyas , M. Bagheri , and T. Phung . "Real-time dry-type transformer aging evaluation." 2017 International Symposium on Electrical Insulating Materials (ISEIM) 2017.

[3] Long, Jiang, et al. "Research on Transformer Fault Diagnosis Based on BP Neural Network Improved by Association Rules." 2019 2nd

International Conference on Electrical Materials and Power Equipment (ICEMPE) 2019.

[4] Beniwal, N. S. , D. K. Dwivedi, and H. O. Gupta. "Life estimation of distribution transformers considering axial fatigue in loose winding conductors." Engineering Failure Analysis 18.1(2011):442-449.

[5] Duan, Ruochen, and F. Wang. "Fault Diagnosis of On-Load Tap-Changer in Converter Transformer Based on Time-Frequency Vibration Analysis." *IEEE Transactions on Industrial Electronics* 63.6(2016):3815-3823.

[6] Jun, Lin, et al. "Prediction Method for Power Transformer Running State Based on LSTM_DBN Network." Energies 11.7(2018):1880.

[7] Matsui, T. , et al. "Development of remaining life assessment for oil-immersed transformer using structured neural networks." Iccassice IEEE, 2009.

[8] Dengwei, Ding, et al. "Research on the winding and iron core operation state of transformer based on the vibration acoustic fingerprint." 2016 International Conference on Condition Monitoring and Diagnosis (CMD) IEEE, 2016.

[9] Hong, Kaixing, et al. "A vibration measurement system for health monitoring of power transformers." Measurement 93(2016):135-147.

[10] Zhou, et al. "Transformer winding fault detection by vibration analysis methods." Applied Acoustics (2016).

[11] Zhao, Shutao, et al. "Intelligence Expert System of Transformer Running State Diagnosis Based on Acoustic Signal Analyzing." International Symposium on Knowledge Acquisition & Modeling IEEE, 2009.

[12] Belén García, Juan Carlos Burgos, ángel Alonso,Winding deformations detection in power transformers by tank vibrations monitoring,Electric Power Systems Research,Volume 74, Issue 1,2005,Pages 129-138.

[13] Bartoletti, C. , et al. "Vibro-acoustic techniques to diagnose power transformers." IEEE Transactions on Power Delivery 19.1(2004):221-229.

[14] Lu, Zhumao, et al. "Measurement and analysis of UHV transformer noise with sound intensity and vibration method." 2017 20th International Conference on Electrical Machines and Systems (ICEMS) 2017.

[15] Lu, Zhumao, et al. "Noise analysis of UHV power transformer and research on active noise reduction system." 2017 20th International Conference on Electrical Machines and Systems (ICEMS) 2017.

[16] Zhang, Pengning, et al. "Study on Vibration of Iron Core of Transformer and Reactor Based on Maxwell Stress and Anisotropic Magnetostriction." IEEE Transactions on Magnetics PP.2(2019):1-5.

[17] Duan, Xiaomu, et al. "Analysis of Winding Vibration Characteristics of Power Transformers Based on the Finite-Element Method." Energies 11.9(2018):2404.

[18] Xiaowen, Wu, et al. "Diagnosis of DC Bias in Power Transformers Using Vibration Feature Extraction and a Pattern Recognition Method." Energies 11.7(2018):1775-.

[19] C. Mu, C. Zhang, X. Huang and Y. Ma, "The efficiency analysis of the statistical feature in network traffic identification based on BP neural network," 2013 5th IEEE International Conference on Broadband Network & Multimedia Technology, Guilin, 2013, pp. 70-74.

The Control Method of Tying Shoelaces for Robotic Hand Based on Angular Velocity Sensor

Zhigang Li
School of Instrument Science and
Engineering
Southeast University
Nanjing, China
220193290@seu.edu.cn

Jianwei Cui*
School of Instrument Science and
Engineering
Southeast University
Nanjing, China
cjw@seu.edu.cn

Pudong Lu
School of Instrument Science and
Engineering
Southeast University
Nanjing, China
220193325@seu.edu.cn

Huice Jiang
School of Instrument Science and
Engineering
Southeast University
Nanjing, China
220193332@seu.edu.cn

Abstract—**Using the human body's signals to control assistive robot has become a focus, but factors such as the way of acquiring signals and the ability to decode signals affect the accuracy of controlling robotic hand. As a wearable device, the angular velocity sensor can directly obtain the movement information of the upper limbs. This paper proposes a new control method for robotic hand based on angular velocity sensors, and this control method is applied to the operation of tying shoelaces with the robotic hand. The angular velocity sensor is used to collect the angular velocity of the human upper limb movement, and the characteristics of fingers during grabbing are analyzed by extracting the characteristic parameters of the mean, the maximum value and the difference value. The experimental results show that the accuracy of controlling robotic hand reaches 99% and the delay time of opening or closing the fingers is less than 1.5s.**

Keywords—angular velocity sensor, robotic hand, tying shoe laces

I. INTRODUCTION

So far, there are about 85 million persons with disabilities in China, of whom about 25 million are physically disabled, and people with hand disabilities account for about one third of the physical disabilities[1].The artificial dexterous hand which is an extension of the limb function of the hand-disabled person, has the advantages of strong versatility, flexible operation, and stable grasping, and plays an increasingly important role in the field of assisting the disabled[2]. Therefore, improving the stability and accuracy of the behavior control of the robotic hand in the completion of daily life movements can help the disabled person achieve self-care in life more effectively.

There are some signals in the human body that can reflect people's will. And by using these signals, the robotic hand can be controlled, mainly including electromyography (EMG), electroencephalography (EEG), and voice control[3],[4],[5]. EMG control relies on electrical signals in the muscles of the limb to complete the control of the robotic hand. But some factors such as the degree of muscle atrophy and the position of the residual limb make the EMG signal in the muscle unable to be accurately collected[6]. Because the EEG signal is extremely weak and mixes with an amount of noise, there are very high requirements for the decoding ability of the EEG

signal. Yuriy Mishchenko et al.[7] designed a three- to six-state electroencephalography based BCI system to control the assistive robot, with an average accuracy rate of 80%.

It can be seen that the way of acquiring signals and the ability to decode signals affect the accuracy of controlling robotic hand. Wearable sensors have a wide range of applications in the direction of arm motion measurement with its advantages of strong wearability, low power consumption and high cost performance[8].Most manipulator control methods need to collect the posture information of the human arm first, and then convert it into the joint variable of the manipulator. Shen Shu et al. [9] used inertial sensors to detect the motion posture of the human arm, and then controlled the mechanical arm to reproduce the movement of the human arm. Feng Lin Er [10] collected the motion postures of the arms and wrists through a nine-axis posture sensor, realizing the real-time movement of the manipulator following the hand.

Taking account of the shortcomings of EMG control and EEG control and taking the signal acquisition and analysis methods as the starting point, this paper proposes a control method for robotic hand based on angular velocity sensors. By analyzing the characteristics of the angular velocity changes of the upper arm, forearm and palm of a healthy person when tying shoelaces, the timing of finger opening or closing can be judged, and the control method can be applied to the control of tying shoelaces tied by the robotic hand.

II. METHOD

A. Sensor Signal Preprocessing

Due to the internal design of the sensor, the shaking of the arm and the drift caused by the sensor working for a long time [11], the real sensor data is mixed with noise interference. Therefore, the Kalman filter is used to filter the datas collected by the angular velocity sensor. The process of Kalman filtering is divided into two parts: prediction and correction. In order to obtain the true angular velocity value at the kth time, in the prediction stage, the Kalman filter predicts the sensor data at the kth time from the sensor data at the k-1th time, and obtains the observation variables at the same time. In the correction phase, the filter corrects the sensor datas obtained in the prediction phase with the observed variables to obtain the optimal estimate at the kth time. Among them, process variance Q=1e-3, measurement variance R=1e-2.

---

* Jianwei Cui is the corresponding author. (e-mail: cjw@seu.edu.cn).

As shown in Figure 1, Figure 1(a) is the raw datas collected by the angular velocity sensor, and the three curves are respectively the x-axis acceleration, the y-axis acceleration, and the z-axis acceleration of the boom. With the increase of sampling points, the datas collected by the angular velocity sensor fluctuate frequently, indicating that noise has a greater impact on the sensor. In Figure 1(b), the angular velocity curves of the three axes are processed by the Kalman filter. Throughout the sampling interval, the angular velocity value changes smoothly, and the fluctuation is significantly reduced, which shows that the effect of random noise is effectively reduced after filtering.
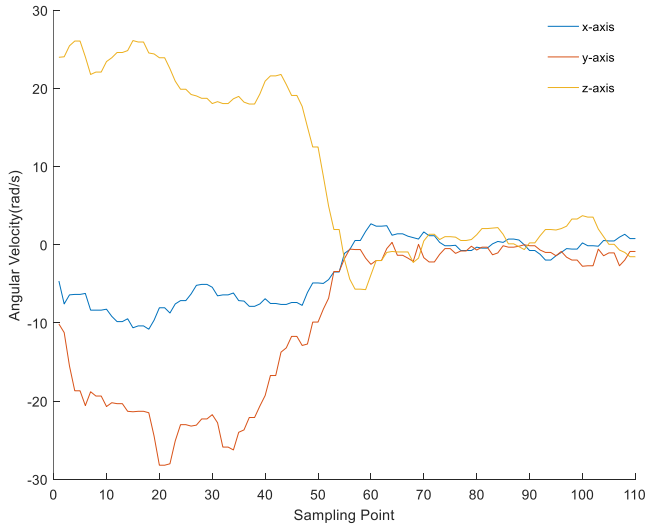
(a) Angular velocity before filtering
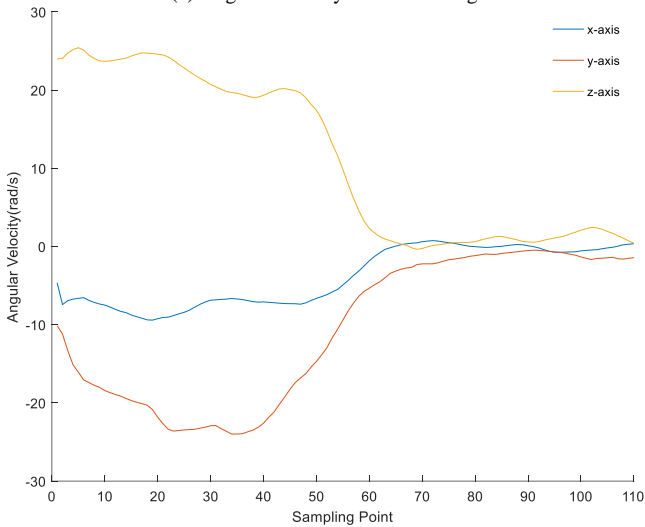
(b) Angular velocity after filtering

Fig 1. Angular velocity curve of boom.

B. Extracting Feature Parameters

In the process of upper limb movement, the angular velocity will change differently as the arm state switches. Therefore, choosing the feature parameters from the angular velocity signals can more accurately represent the behavior characteristics of human body. Combined with the analysis of the changing law of angular velocity, this paper uses the time domain method to analyze the sensor signal, and extracts the following characteristic parameters.

1. Mean

Mean of angular velocity is calculated as

$$\overline{\omega} = \frac{1}{N} \sum_{i=1}^{N} \omega_i \qquad (1)$$

where N is the number of sampling points and ω_i represents the angular velocity value of sampling point i.

2. Maximum

In the process of upper limb movement, the two sides of the maximum angular velocity datas correspond to the opposite trend of angular velocity change.

3. Difference

The difference value of two adjacent sampling points can indicate the change speed of angular velocity. The larger the absolute value of difference, the faster the change of angular velocity. The symbol of difference value indicates the change direction of angular velocity. Difference of angular velocity is calculated as

$$\Delta \omega_i = \omega_i - \omega_{i-1} \qquad (2)$$

C. Robotic Hand Control Scheme

Normally, in order to grasp objects stably, healthy people will open or close their fingers when their arms are at rest. Based on the analysis of the datas measured during the upper limb movement, it is found that when the arm moves from motion to rest, the angular velocity of the upper limb fluctuates greatly, and the value of the angular velocity tends to zero. When the angular velocity of the upper limb drops to zero, it indicates that the arm is at rest. Therefore, when studying the motion law of the arm, the opening or closing of the fingers of the robotic hand can be controlled from the changing characteristics of the angular velocity of the upper limb.

The key to complete the action of tying shoelaces is to grasp the timing of grabbing the shoelaces, that is to open or close the fingers to control the shoelaces in time. Figure 2 records the change curve of the angular velocity of the forearm when the upper limb grabs the object. The horizontal axis represents the sampling point, and the vertical axis represents the angular velocity. The blue curve, red curve and yellow curve are sensor datas in the x-axis, Y-axis and z-axis directions respectively. The arm motion in three-dimensional space is decomposed into three axes. The x-axis represents left and right space, the y-axis represents front and back space, and the z-axis represents up and down space.
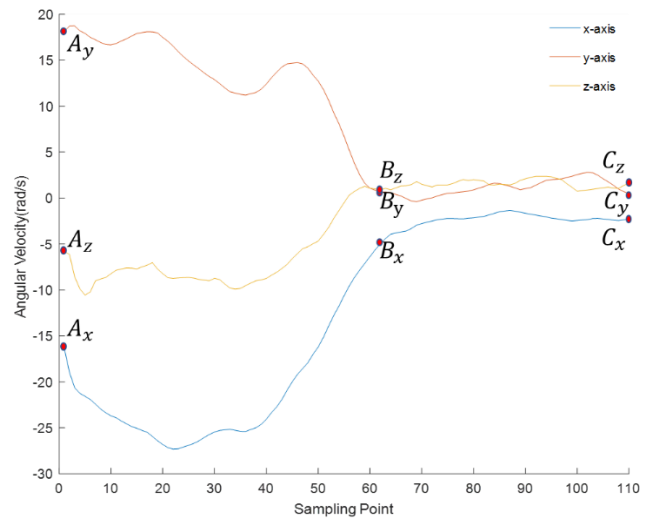

Fig 2. Angular velocity curve of forearm

In the curve diagram of the forearm angular velocity, segment AB represents the angular velocity of the forearm during motion, and segment BC represents the angular velocity when the forearm is stationary. In the upper limb movement, the angular velocity is greater than 5, and there is a large range of fluctuations. When approaching point B, the angular velocity of three axes tends to zero obviously, and the symbol of difference of several continuous sampling points is the same. When the forearm is stationary, the angular velocity of Y-axis and z-axis are obviously less than 5, and fluctuate near zero. In the process of upper limb movement, the change trend of angular velocity is consistent with the kinematic law. Therefore, according to the change characteristics of the maximum and difference of angular velocity, the movement state of upper limb can be distinguished to control the opening and closing of the robotic hand.

III. EXPERIMENTS SET-UP

The content of the experiment is that the healthy person holds the assistive hand to complete the action of tying the shoelace. Five healthy volunteers were involved in the experiment. Each volunteer wore the experimental equipment and completed the task of tying shoelaces according to the experimental requirements. Each person did 10 experiments, and a total of 50 experimental datas were obtained.

A. Experimental Devices

The experimental equipment consists of two parts: the data glove and the robotic hand, as shown in Figure 3. The wearable data glove device integrates mpu6050 inertial sensor, which is distributed in the palm, forearm and upper arm of the upper limb, and can output angular velocity datas. The robotic hand is connected with the single chip microcomputer. When the single chip microcomputer receives the control command from the PC terminal, it immediately outputs a level to control the finger of the robotic hand to open or close. The angular velocity of the upper limbs is collected by the data glove. The PC terminal is responsible for processing the datas and sending control command to the single chip microcomputer.
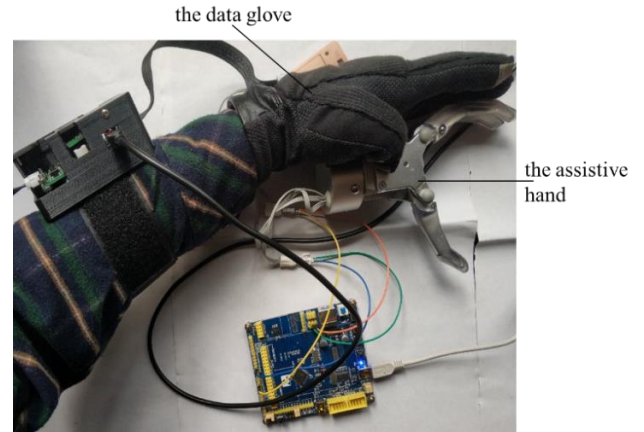
Fig 3. Experiment equipment

B. Decomposition Movement of Tying Shoelaces

This article chooses a simple action of tying shoelaces, which is easy to operate. In order to accurately study the law of upper limb movements when tying shoelaces, the selected action is decomposed in accordance with the sequence of the movements, and each decomposition action is bounded by the opening or closing of the fingers. The schematic diagram includes the unidirectional movement, the flipping, the moving back and forth and the twitch, as shown in Table 1. The circle represents the elbow, the rectangle represents the finger. The dotted line represents the state before the upper limb moves, and the solid line shape represents the state after the upper limb moves.

TABLE 1. SCHEMATIC DIAGRAM OF DECOMPOSITION

Action category	Demonstration of action	Description of action
Unidirectional movement		From the starting position, the upper limb moves in a certain direction to the target position for grabbing.
Flipping		When the finger grasps the target, the elbow is used as the fixed point, and the palm is turned in a certain direction.
Moving back and forth		After moving the upper limb back and forth from the preparation position, it still reaches the preparation position.
Twitch		After grasping the target, the finger moves quickly in a certain direction.

IV. RESULTS AND DISCUSSION

This article evaluates the feasibility of the control method from the following two aspects.

A. The Accuracy of Controlling the Robotic Hand

It is the key to tie shoelaces to accurately control the opening or closing of the fingers. Since the change trend of the angular velocity of the forearm, forearm and palm is similar in the process of tying shoelaces, the angular velocity of the forearm of a sample is selected for analysis. Consider that after completing the movements of the unidirectional movement, the flipping, the moving back and forth and the twitch, it is necessary to repeat the action to tie shoelaces, which generates plenty of datas, so only the datas of the first movement is shown in Figure 4.

In Fig. 4, segment AB is the preparatory stage before the upper limb moves, at which stage the upper limb is stationary. The CD segment, EF segment, GH segment, IJ segment and KL segment were all in static stage after upper limb movement. At this stage, the angular velocity of the forearm is close to 0, and the angular velocity changes stably. Five volunteers conducted a total of 50 experiments, each of which can accurately control the opening and closing of the fingers, and the accuracy of controlling robotic hand reaches 99%. The experimental results show that the same symbol of the difference value of a group of sampling points is used as a sign that the upper limb has moved, and the maximum value of a group of sampling points is less than $\pm 5 rad / s$ as the sign of upper limb showing static state. Combined with these two marks, the opening and closing of the fingers of the robotic hand can be accurately controlled. The control effect of the robotic hand is shown in Figure 5.

The changes in angular velocity can be used to distinguish different movements when tying shoelaces. The BC segment is the unidirectional movement of the upper limbs. Compared with other stages, the angular velocity datas change slowly at this time. The DE segment and HI are the flipping, the angular velocity datas show a clear single peak or trough, and the number of sampling points in this stage is relatively small. The FG segment and the JK segment are the moving back and forth,

and peaks and troughs will appear continuously at this time. The LM segment is the twitch. Compared with the flipping, the amplitude of the peak or trough appearing in the twitch is relatively small. More importantly, for any axis, if a peak appears in the flipping, a trough appears in the twitch. This is because the direction of flipping and twitch are opposite.

B. The Delay of Finger Movement

Low latency is another key factor in judging whether the robotic control method is feasible. It is stipulated that the delay time of each finger movement is less than 2s. Therefore, the time point of upper limb rest and the time point of finger opening or closing are recorded, and the difference between the two time points is used as the degree of delay in judging the movement of finger. Considering the large amount of datas in the experiment samples, the record results of one experiment were randomly extracted from each volunteer's ten experiments. In the extracted record results, only the delay time of the first action is selected, as shown in Table 2. The $t_1 \sim t_4$ in turn is the difference between the two time points in the opening or closing of fingers of the robotic hand in the first movement.

TABLE 2. DELAY TIME OF FINGER MOVEMENT

Number	$t_1(s)$	$t_2(s)$	$t_3(s)$	$t_4(s)$
1	1.31	1.23	1.32	1.36
2	1.34	1.32	1.36	1.39
3	1.32	1.36	1.33	1.41
4	1.30	1.29	1.35	1.47
5	1.27	1.32	1.32	1.37

It can be found from Table 2 that, in the extracted record results, the delay time of opening and closing the fingers is less than 2s, which is in accordance with the prescribed delay time of finger movements.
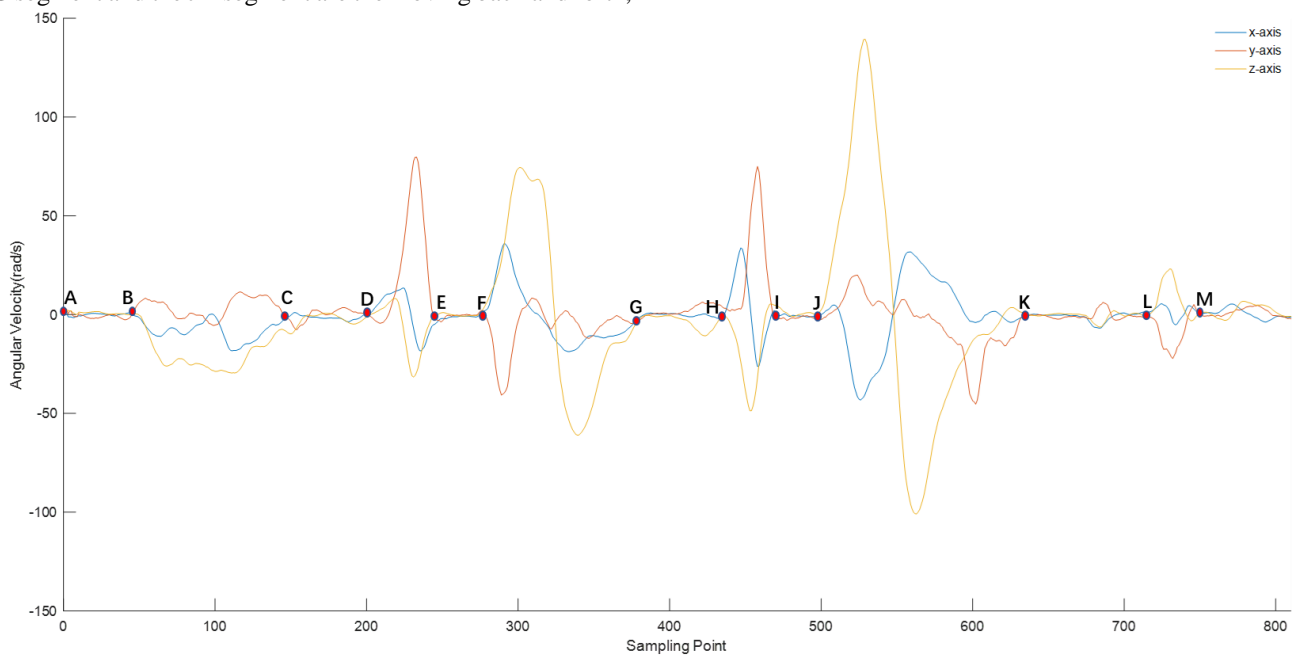

Fig 4. Curve of the angular velocity of the forearm when tying shoes

Fig 5. The operation of tying shoelaces

V. CONCLUSION

This paper presents a method for controlling the robotic hand based on the angular velocity sensor. Using angular velocity sensors to collect angular velocity datas during human upper limb movements, the motion characteristics of fingers movement were studied by preprocessing sensor signals and extracting characteristic parameters. Finally, the control method is applied to the experiment of tying shoelaces with robotic hand. The experimental datas were analyzed from the accuracy of controlling the robotic hand and the delay of finger opening or closing. The next work will focus on the flexible control of the robotic hand, such as the realization of the rubbing control between the fingers, rather than the mechanical closing of the fingers. Reducing the delay time of finger movement is another important content of future work.

REFERENCES

[1] http://www.cdpf.org.cn/sjzx/cjrgk/201206/t20120626_387581.shtml.

[2] V.N.Iliukhin, K.B. Mitkovskii, D.A. Bizyanova, "The modeling of Inverse Kinematics for 5 DOF manipulator", Elsevier, Feb 2017, pp.498-505.

[3] Z.Z.Luo, W.G.Li, "A design of two -freedom EMG artificial hand with real -time control", Chinese Journal of Rehabilitation Medicine, vol. 24, no. 4, pp. 355-358, Apr. 2009.

[4] B. Dany, N. Amrita, S. Devashish, B. Prof.Swati, "EEG-Based Brain Controlled Prosthetic Arm", IEEE, 2016 Conference on Advances in Signal Processing, Jun 2016, pp.479-483.

[5] J. Vikram, S. D. Hardeep, B. Sandeep. "Radio Frequency Home Appliance Control based on Head Tracking and Voice Control for Disabled Person", IEEE, Fifth International Conference on Communication Systems and Network Technologies, Fab 2015, pp.559-563.

[6] G. L. Li, W. S. Oluwarotimi, L.Chuang, G.A.Mojisola, F.Peng, O.I.Paul, "Realizing Efficient EMG-Based Prosthetic Control Strategy", Springer, Nov 2019, pp.149-166.

[7] M. Yuriy, K. Murat, O. Erkan, Y. Hilmi, "Developing a Three- to Six-state EEG-based brain-computer interface for a virtual robotic manipulator control", IEEE Transactions on Biomedical Engineering, Apr 2019, pp.977-987.

[8] X. M. Wang, Z. Liu, W. B. Ni, "Motion measurement and recognition method of human arm based on MEMS sensors", Journal of Chinese Inertial Technology, vol. 25, no 6, pp.701-707, Dec 2017.

[9] S. Shu, K. Gu, X. Y. Liu, "Design of an Intelligent Wearable System for Physical Disability", Computer Technology and Development, vol. 29, no.7, pp.124-129, Jul 2019.

[10] L. E. Feng, "Research on Attitude Synchronization and Force Feedback Design and Algorithm of Humanoid Manipulator", University of Electronic Science and Technology of China, 2019.

[11] Kwapisz J R, Weiss G M, Moore S A, "Activity Recognition Using Cell Phone Accelerometers", ACM Sig KDD Explorations Newsletter, Jul 2011, pp.74-82.

Design and Performance Analysis of a Moveable Groove-tracking Milling Machine for Repairing Thread of Marine Steering Gear Nuts

Guiping Lu*
Beijing Institute of technology, Zhuhai
Guangdong China
lgpcan@163.com

Zhiyong Xiao
Beijing Institute of technology, Zhuhai
Guangdong China
xiaozhiyong96@163.com

Zhensheng Zhong
Beijing Institute of technology, Zhuhai
Guangdong China
AaronZhongcn@outlook.com

Jiaran Liang
Beijing Institute of technology, Zhuhai
Guangdong China
jiaran_liang@163.com

Zhili Xiang
Beijing Institute of technology, Zhuhai
Guangdong China
catzhili@163.com

Zhiyang Tang
Beijing Institute of technology, Zhuhai
Guangdong China
857317510@qq.com

Abstract—In the shipping industry, parts of a ship are prone to damage after a long voyage. Due to working under high pressure and high load for a long time, the connecting thread of the nut will get stick, which makes the parts in poor operation. This failure frequently occurs at the nut connection between tiller and rudderstock, as well as between the propeller and the stern shaft. The nuts with damaged threads are often repaired by a method which called "groove-tracking machining". However, due to the huge volume of the nut, generally 300-1,500mm in nominal diameter, it takes lots of time to disassemble, assemble and transport it to the workshop. And there are problems such as finding the origin location of the screw helix, which leads to high maintenance cost for the large nuts. Aiming at repairing damaged internal thread of the nuts for marine steering gear, this paper presents a design of a groove-tracking milling machine, which is moveable and easy to assemble and disassemble. Due to the huge volume of the nut, this design will adopt a structure, which makes the nut fixed and the whole machine built into the nut to repair the thread. Further-more, finite element analysis, kinematics simulation analysis and dynamics analysis are carried out on the main parts by UG NX, which provide necessary reference and theoretical basis for the production and manufacture of this machine.

Keywords—*Marine nuts, Repair of large internal thread, Four-axis feed system in cylindrical coordinate system, Finite element analysis.*

I. INTRODUCTION

Aiming at repairing for the huge nuts used in marine steering gear, this paper develops a kind of automatic machine which has the advantages of simple control, high processing precision, and can carry out groove-tracking machining on the spot to reduce labor cost and transportation cost. The groove-tracking machining is used to remove the sintered layer from thread surface and then grind it with a whetstone until it is smooth. According to functions, the mechanical structure systems are divided into three parts, including spindle system, four axis feed system and support system. In order to achieve moveable, each part of the system should be designed as a relatively independent module, and thus to realize the repairing process of "disassembly-transport-assemble-machining" on the spot. Among them, the spindle system is used to drive the tool to process the thread along its helical cutting path. This system is connected to the four-axis feed system by a flange connection, thereby implementing modularity. Therefore, the appropriate principal axis can be replaced in different working conditions, so that the process capability is improved, and the stiffness and kinematic accuracy of the system can be improved.

In this design, we carried out a detailed design for the mechanism and established the 3D model by UG NX. At the same time, we performed finite element analysis, motion simulation analysis and dynamic analysis of the main components. Combining with the feed control method, the motion trajectory of each part was obtained. Then, we analyzed the deformation by applying force to components and compared the theoretical trajectory with the actual trajectory for error analysis.

II. SYSTEM DESIGN OF THE MACHINE

A. Design Flow

The groove-tracking machining method is used to remove the sintered layer from thread surface, and its main motion is the rotation of the spindle, while the feed motion is the spiral motion of the spindle driven by the machine.

The first step is to calculate the cutting force when milling thread along the helix groove, and then based on it to design the structure and size of the spindle. Secondly, due to the design speed and other requirements, the servo motor is selected as the spindle motor. Then through the selection of standard parts such as motors and bearings, based on their size, the structure and size of the spindle box can be designed. This method is also applicable to the design of the feed mechanism. When designing the structure, the processing technology of parts should be considered. After the design, to conduct FEA on the parts to analyze the deformation, and thus to determine whether the parts meet stiffness requirements [1-4].

B. Overall Design Scheme

According to the processing characteristics of the thread, this machine must meet the following three performance requirements:

1) It can move and rotate simultaneously in the vertical direction and the rotation direction: By controlling the speed ratio between the two movements, helixes with different pitches are obtained for groove-tracking.

2) It can move in the horizontal direction: By changing the distance between end effector and the centerline of the middle shaft, it can match nominal diameters of threads that need to be repaired.

3) The end effector(the spindle head): it can rotate based on angular positioning to meet the processing of both sides of thread profile and the processing of different types of threads.

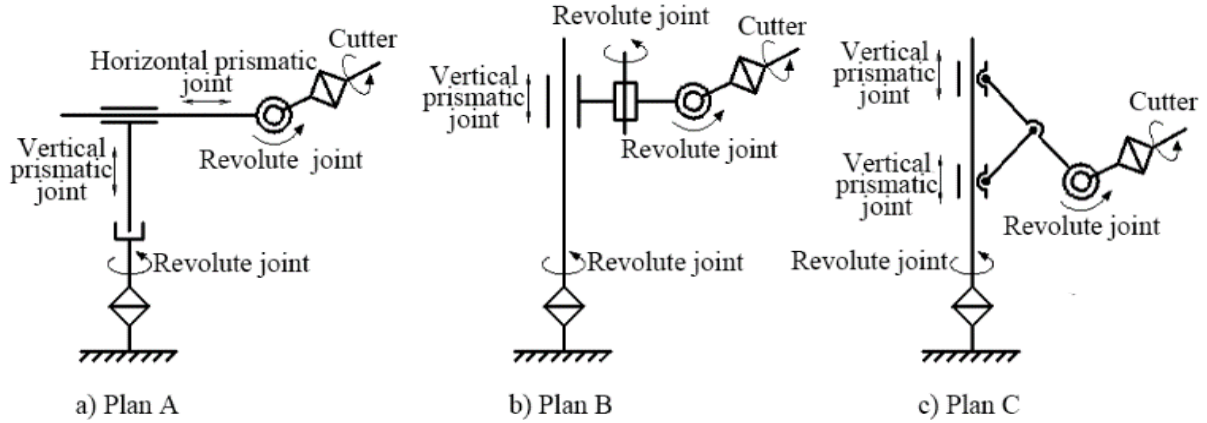Based on three requirements, the following three design schemes are proposed, as shown in FIG.1:

Fig. 1. Schematic diagram of feed mechanism of the groove-tracking milling machine.

- Plan A: It has the same structure as the RPP cylindrical coordinate robot, with two prismatic joints (2P) and one revolute joint (1R). Based on it, the groove-tracking milling machine adds one more revolute joint and a spindle head at the end of the horizontal prismatic joint.

- Plan B: This structure is derived from the structure of SCARA robot, the difference is that it adjusts the vertical prismatic joint from the original position at the end to above the revolute joint at the center axis [5].

- Plan C: A new type of linkage mechanism, which drives the connecting rod through the movement of two vertical prismatic joints, so that the end effector can extend or retract. The structure of vertical prismatic joint is the same as that of plan B, both of which are rotating nut ball screws. In this plan, the two joints are driven by two motors separately. And combined with the base motor driving the screw rod and the vertical shaft guide rail to rotate, the end effector can move horizontally and vertically, and rotated horizontally and spirally. At the same time, the angular positioning of the spindle head can be achieved by driving its revolute joint.

Compared with the three plans, plan C is more reasonable. The following reasons:

1) It has the characteristics of being collapsible and retractable in the axial direction.

2) Its structure has the characteristics of special geometric: The triangle formed by the two rods is an isosceles triangle, the length of short rod is half of the length of long rod, and the joint of two rods is in the middle of the long rod, which makes the design and control easier.

3) This plan has the characteristics of stroke integrity in the vertical dimension: In geometric, it can be deduced that the position of the end revolute joint is aligned with the connection joint of short rod on a horizontal plane at any position, so the end effector will never interfere with the bottom surface at lowest stroke position in vertical direction.

C. Calculation of Cutting Force

According to selection of quantities in cutting, the processing technology is a precision machining, which has the

characteristics of small feed rate and high feed speed. Due to the uncertainty of thread profile and pitch of the nut, it will be set

as a larger coarse thread when calculating the cutting force, to meet the requirements of a serrated nut with small pitch.

By NX, we enlarged the ordinary hex nut through modeling to obtain a maximum hypothetical nut, and then used it as the standard for calculating the maximum cutting force. It can be known that the machine uses a milling method to machine the thread surface. Similarly, taking a common triangular thread as an example, we designed 3-flute tapered end mill, as shown in Fig. 2:
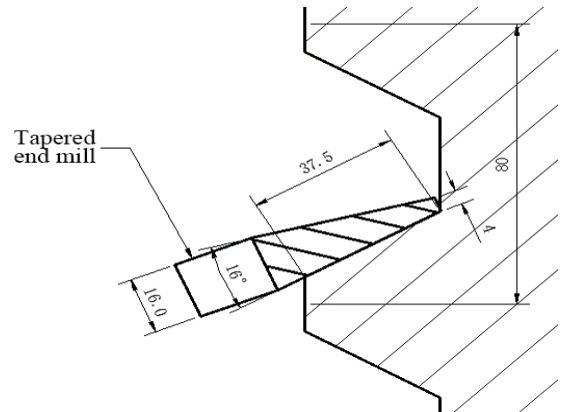
Fig.2. Using 3-flute tapered end mill to machine screw thread along the thread profile.

Setting: the material of the nut is divided into steel and cast iron. The milling cutter diameter d_0=16mm, the rake angle γ_o = 0°, the helix angle ω = 30°, the depth of cut a_p = 0.05mm. The empirical formula for calculation of the cutting force is:

$$F_c = \frac{c_p a_p^{0.86} f_z^{0.74} Bz}{d_0^{0.86}} \times K \times K_1 \times 10 \qquad (1)$$

Where: B(mm)-the milling width, z-the number of cutting edge, c_p-the influence coefficient of workpiece material on milling force, K-the influence coefficient of cutter front angle on milling force, K_1-the influence coefficient of milling speed on milling force, f_z(mm/z)-the feed engagement, d_0(mm) -the milling cutter diameter.

D. Structural Design of the Manipulator

The structure of the connecting rod manipulator of the feed system is an isosceles triangle. Based on it, set the length of the long rod L_1=500mm, and the length of the short rod L_2=250mm. The joint where the two links intersect is the midpoint of the

long rod, so the schematic diagram of the connecting rods can be drawn, as shown in Fig. 3:
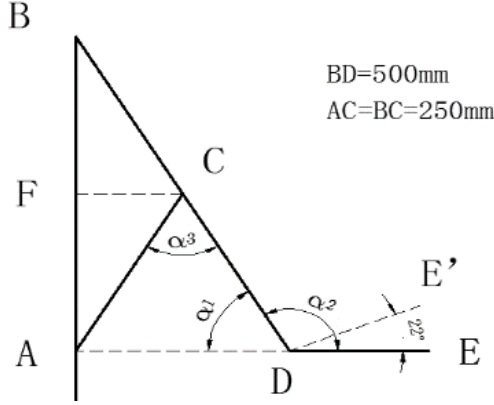

BD=500mm
AC=BC=250mm

Fig.3. Schematic diagram of the connecting rods.

The long rod is the upper mechanical arm. Through mechanical analysis, we find that it is subjected to tension and bending moments under no-load condition. Therefore, in terms of the cross-sectional shape of the arm, a hollow rectangular structure is adopted to achieve weight reduction and attractive appearance. After determining the size of the spindle head and the motor selection of the revolute joint, the end size of the mechanical arm is basically determined. The specific three-dimensional structure of the upper mechanical arm is shown in Fig. 4:
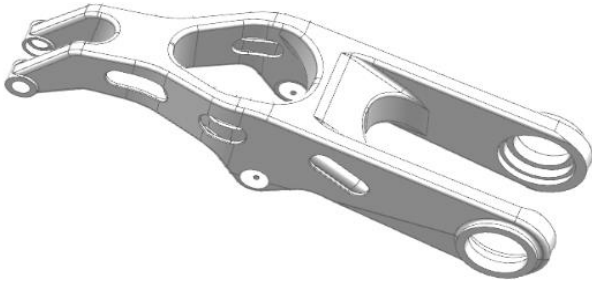
Fig.4. Three-dimensional structure of the upper mechanical arm.

E. Selection of Motor Used in Base of Middle Shaft

As the main rotary motion of the entire feed system, the drive power required by the motor of middle shaft is the largest. According to the force analysis during groove-tracking cutting, the middle shaft is mainly subject to the feed force of rotary cutting and the tensile force of the robot arm, and the maximum torque is 189 Nm. According to the virtual assembly analysis by UG, the moment of inertia of the manipulator relative to the center axis at the maximum extended posture can be calculated: $I_{Zmax} = 4.94$ kg/m².

According to the calculation of cutting parameters, it can obtain that the maximum feed rate of cast iron $V_f = 1800$ mm/min, and the minimum inner diameter $d_{min} = 500$ mm. Also, it can be roughly calculated that the maximum rotation speed of the middle shaft n1 = 1r/min during cutting. To reduce the time of no-load return and improve efficiency, set the following parameters: the maximum rotating speed of middle shaft n2 = 60 r/min; the accelerated speed a = 10π/s². According to the torque formula, the starting torque can be obtained: $M_{start} = I_{Zmax} \cdot a = 155$ Nm.

Based on the data obtained from the preliminary calculation and the information provided by the motor supplier, the motor model was selected as SHA40A80CG, and its technical parameters are shown in Fig. 5:
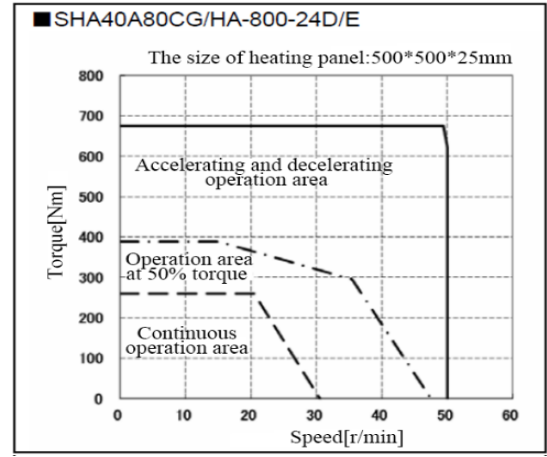
Fig.5. Torque-speed curve illustrating operation area of electric motor.

The calculation the maximum load moment is shown in equation (2), and the schematic diagram is shown in Fig.6. It can be known the requirement that: the maximum load moment (Mmax)< allowable moment (Mc).
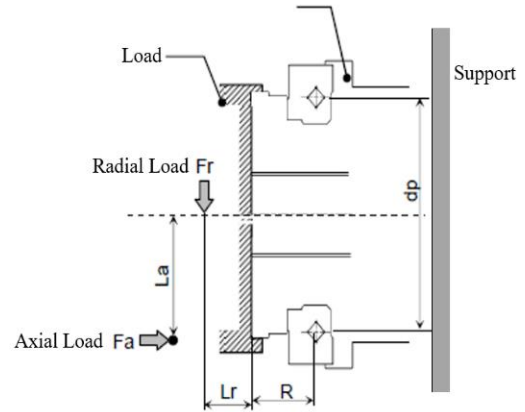
Fig.6. The maximum load moment.

The calculation the maximum load moment is shown in equation (2), and the schematic diagram is shown in Fig.6. It can be known the requirement that: the maximum load moment (M_{max})< allowable moment (M_c).

$$M_{max} = \frac{Fr_{max} \cdot (Lr + R) + Fa_{max} \cdot La}{1000} \qquad (2)$$

Where: M_{max}- the maximum load moment, Fr_{max}- the maximum radial load, Fa_{max} - the maximum axial load, L_r and L_a- the moment arm, R- the amount of offset.

Check of allowable moment of motor:

According to the definition in Fig.6, the data measured in the assembly drawing is plugged into the formula to obtain that $M_{max} = 134.7$ Nm, which is less than the allowable moment Mc(849Nm). Therefore, the middle shaft can be directly connected to the motor through the flange seat. According to the definition of rotational stiffness, when the upper end of the vertical shaft is a free end, its inclination angle θ can be obtained:

$$\theta = \frac{134.7}{179 \times 10^4} \times \frac{180^0}{\pi} = 0.00431^0 \qquad (3)$$

Since the upper support structure of the vertical shaft is supported by a crossed roller bearing, the resultant of inclination

angle θ and the deformation of the vertical shaft still meets the rigidity requirements [7].

F. Design of the Base Chuck

The core structure of the base clamp is a linkage mechanism of three jaws, its linkage system is composed of three screws connected to three small bevel gears and a large bevel gear as an idler. In this way, by rotating one of the screws, it can move the three jaws in unison. And the jaws can slide on the base along the triangular sliding guide.

When clamping, the nut is lifted and suspended above the center of the chuck. Driven by three claws, the nut is moved so that its axis gradually approaches the axis of the column, and then the nut is gently released after being tightened. The jaws have rollers on the sides, with which the heavy nut can slide up and down to the positioning surface. Due to the large muss of the huge nut, when the cutting force is not large, the static friction can be used to meet the fixing requirements. Combined with the additional clamping force of the jaws, the nut can be clamped tightly. The structure of the three-screw linkage system is shown in Fig. 7:
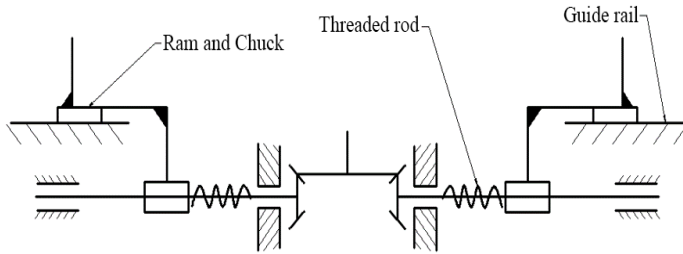
Fig.7. Three-screw linkage system.

G. Design of Upper Support Fixture of Middle Shaft

As shown in Fig. 8, the main structure of upper supporting fixture is composed of threaded rod, connecting rod, supporting arm and chuck. Its working principle is: the rotation of the threaded rod drives the three connecting rod ends to move up and down, to achieve clamping and loosening of the chuck. Through the guide rail of the supporting arm, the magnetic movable chuck can slide on it to adjust the position.

The connection between the support and the vertical shaft is shown in Fig.8. The sleeve bolt of support is a steel pipe with three slots cut, which can be connected to the support as an integral part by welding, and a nylon bushing is affixed in the pipe.
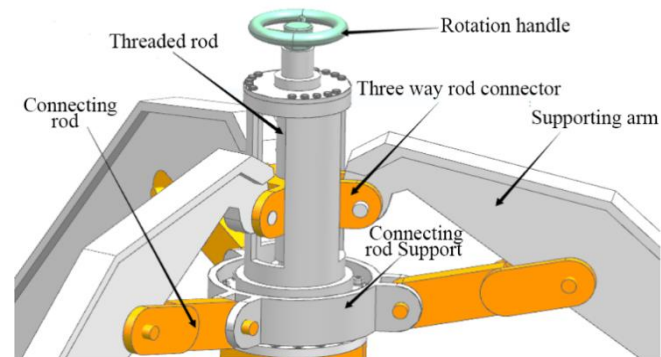
Fig.8. Three-dimensional model of the main structure at the top of the vertical shaft.

H. 3D Modeling of the Whole Machine

Based on Unigraphics NX, the modeling method of this design is to directly mod-el the components in assembly context, so that there is no interference between the components, which

is an efficient method. Through a series of design calculations and part selection, the whole machine modeling and assembly is gradually completed. The overall three-dimensional assembly diagrams are shown in Fig.9 and Fig.10:
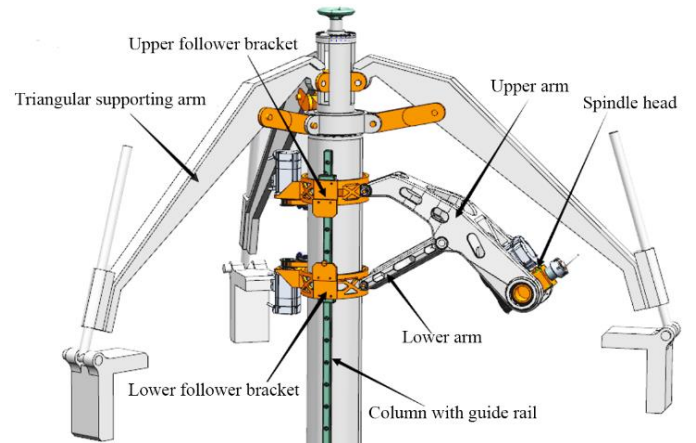
Fig.9. Assembly diagram of the upper part of the machine.
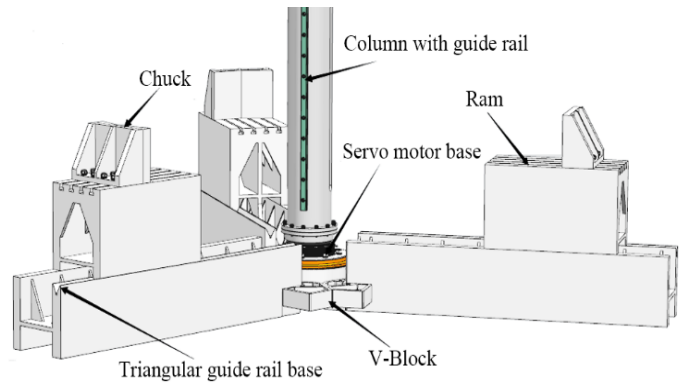
Fig.10. Assembly diagram of the lower part of the machine.

III. FINITE ELEMENT ANALYSIS

A. Check for the Upper Mechanical Arm

The upper mechanical arm is subject to tension and bending moment. Since the point of force is not on the mechanical arm, it should be converted into a torque to act on the output flange by calculation. According to the working condition of the upper arm under the maximum force, we perform a force analysis on it, as shown in Fig.11:
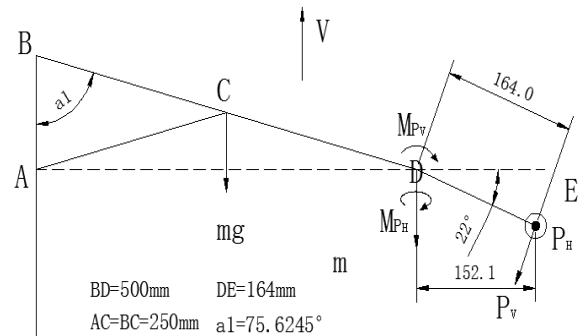

Fig.11. Force analysis of the upper arm.

Through the measure tool in NX, the total mass of the spindle head can be measured that $M_{spindle\ head} = 8.8kg$. According to the force of the milling cutter, the torque received by the upper arm in the case of down milling can be obtained:

$$M_{P_V} = P_V \cdot 164 = 28372\ N \cdot mm \tag{4}$$

$$M_{P_v} = P_H \cdot 152.1 = 39546 \ N \cdot mm \qquad (5)$$

HT200 material properties were assigned to the upper arm, the specific setting of the properties is shown in the Table 1.

TABLE I.   MATERIAL PROPERTIES FOR THE UPPER ARM

Properties (unit)	Density (g/cm3)	Young's Modulus (N/mm2)	Poisson ratio	Shear modulus (N/mm2)	Yield Strength (N/mm2)
Value	7.2	130000	0.25	45000	220
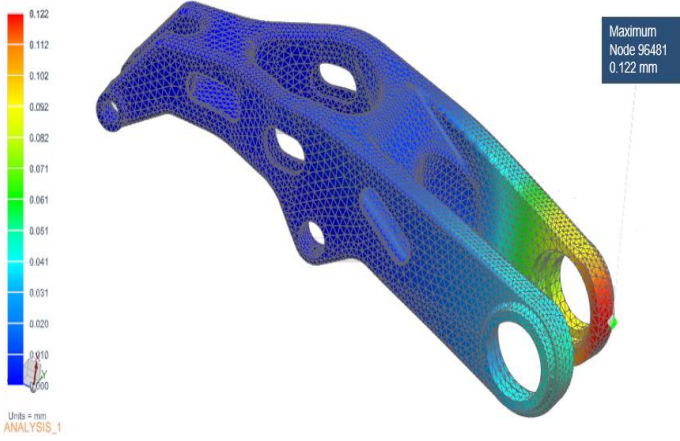
Fig. 12. Displacement nephogram of upper arm in the x-y-z direction.

Fig.13. Stress nephogram of upper arm in the x-y-z direction.

The maximum deformation displacement of the upper arm is 0.122mm, and the direction is the lower inner side，as shown in Fig.12 and Fig.13. Since the spindle head is installed at the shifting fork of the upper arm, the displacement amount on the side will be relatively reduced. For this deformation, it is necessary to adjust the feed amount to compensate according to the actual deformation, and to optimize the structure. In terms of stress, the maximum value is about 17Mpa, which is less than the allowable stress of the material, so it meets the requirements of strength. However, it still needs to increase the stiffness of the upper arm.

B. Analysis of the Whole Assembly

The relationship of the applied load force is shown in Fig.14:

Fig.14. The simulation of applied constraints, connections, and loads.

Apply constraints and loads: To simplify the analysis model, suppose the up-per arm is stationary relative to the absolute coordinate system while the nut is rotated. Based on it, four fixed restraints can be applied to the contact surface of pin bearing on the upper arm. In addition, since the spindle head is connected at the shifting fork, an anti-symmetrical constraint should be applied to the end faces inside the bearing holes on both sides of the shifting fork. [8-10]

After simulation, the result contours can be obtained, as shown in Fig.12 and Fig.13.

Based on the assembly relationship, " Surface to Surface Gluing " object type is applied to the contact surfaces at the connections between the components. On the addition of constraints, the two ends of the vertical shaft and the threaded rod are fixed. The load force is mainly applied in the following forms: the weight of two servo motors of the follower bracket, the weight of the servo motors of the spindle head, and the cutting force.

After simulation, we can get Fig.15, Fig.16 and Fig.17. It can be known that the maximum deformation of the screw is merely about 0.02mm, which has little effect. The place where the overall deformation is the part of cutting edge. And the whole robot arm is deformed downward along the direction of feeding force, of which the maximum displacement is 1.254mm. When processing the nut with the largest diameter, the amount of deformation is too large, that it still needs to improve the connection structure of the mechanical arm and the vertical shaft, to increase the rigidity.
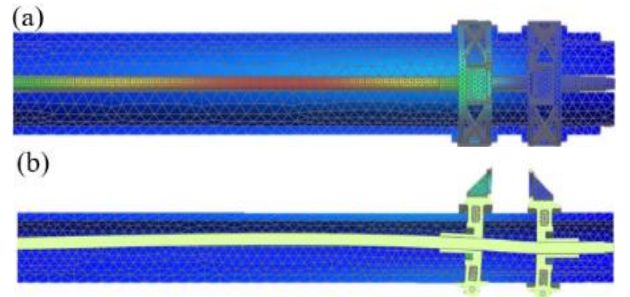

Fig.15. The deformation of the threaded rod: (a) the main deformed part, (b) sectional view.
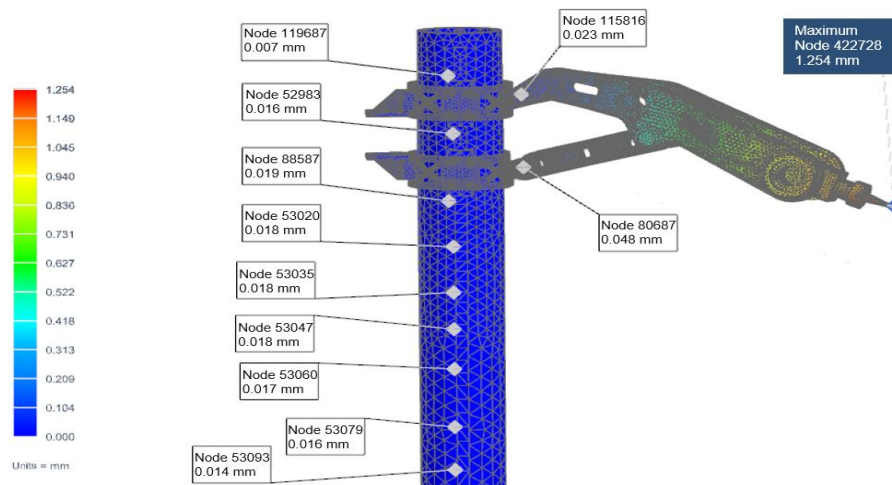
Fig. 16. The displacement contour of the nodes of whole assembly in the x-y-z direction.
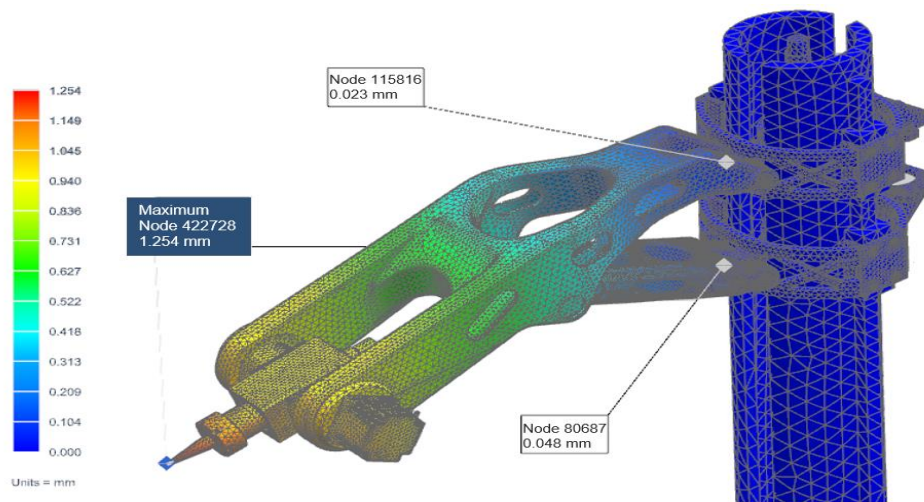

Fig. 17. The displacement contour of the nodes of whole assembly in the x-y-z direction.

IV. CONCLUSION

Based on field investigation of the repairing technology for large nut threads in shipyard, the design scheme of moveable groove-tracking milling machine was put forward. Firstly, we calculated and analyzed the cutting force of the spindle cutter that directly participates in groove-tracking milling for threads. Then we gradually elaborated the structure of the parts connected to it, and finally completed the design of the entire machine, providing a certain basis for subsequent development. Based on advanced simulation in NX, we carried out FEA for the main parts and motion simulation for the whole machine, which further verifies the feasibility of the design.

REFERENCES

[1] W.C. Wang, "Study of Repairing Screw Thread Technology Based on Open NC System," Shenyang University of Technology, March 2006.

[2] F.W. Wan, "The Research on Repairing Technology of Drill Pipe Thread," China University of Petroleum (East China), May 2014.

[3] N.Y. Wang, Y. Shen, "Repairing technology for thread with large size," vol.3, Shanghai:Equipment Machinery, 2014, pp.63-67.

[4] Z.C. He, "Reliable Strengthening repair method for internal thread," , vol. 11, Shanghai: Metal Working, 2000, pp.26-27.

[5] T.S. Liu, "The Key Techniques Research on Industrial Manipulator with PRP Cylindrical Coordinate," Shenyang University of Technology, Februaray 2013.

[6] J.F. Chen, Compendium of metal cutting calculation manual, Jiangsu: Phoenix Science Press, 2007.

[7] X. Xiao, J.T. Yang, F.C. Quan, J. Li, X.B. Tian, P.F. Wang, "Servo Motor Selection Principle and Calculation," vol. 22 ,Guangzhou: Machine Tool & Hydraulics, 2014, pp. 44-46+49.

[8] L. Zhao, "Study on Technology of Milli ng Force Simulation for Virtual Manufacturing," Zhejiang University, January 2005.

[9] H.W. Zhang, E.M. He, Z.B. Zhao, "The Studying of Cross-section Moment of Inertia in the Secondary Development Based on UG," vol. 34, Beijing: Science Technology and Engineering, 2011, pp. 8567-8570.

[10] B. Gosowsk, "Non-Uniform Torsion of Thin-Walled Members of Stiffened Open Steel Structures," vol. 02 , Beijing: Steel Construction, 2008, pp.85.

Research on the Method of Rub-Impact Fault Recognition Based on the Conditional Generative Adversarial Nets and Acoustic Emission

Jing Li
National Engineering Research Center of Turbogenerator Vibration
Southeast University
Jiangsu,China
230149053@seu.edu.cn

Aidong Deng[*]
National Engineering Research Center of Turbogenerator Vibration
Southeast University
Jiangsu, China
dnh@seu.edu.cn

Yong Yang
School of Information Engineering
Southeast University,
Jiangsu, China
yongyang@cumt.edu.cn

Jing Zhu
National Engineering Research Center of Turbogenerator Vibration
Southeast University
Jiangsu, China
1907175086@qq.com

Minqiang Deng
National Engineering Research Center of Turbogenerator Vibration
Southeast University
Jiangsu, China
928188643@qq.com

Abstract—**A large number of effective annotated data is the key support for learning a fault diagnosis model of mechanical equipment. However, the existing practical samples used for training fault classifiers are usually small and interfered by noise .According to this problem the paper presents a rub impact fault recognition method based on the Conditional Generative Adversarial Nets (CGAN) and Acoustic Emission (AE) technology. The data are from the Wind Turbine Train test bed. AE features are extracted from various views such as time, frequency and energy intensity under different operation state. The proposed CGAN model adds useful auxiliary information into each layer of GAN generation model to improve the quality of generated pseudo-samples. It is further to evaluate the probability of samples from the training set or real set.The experimental results show that this model can effectively identify the rub impact fault and have strong robustness. It is an effective way to solve the problem of inadequate sample and improve the recognition performance of rub impact fault.**

Keywords—Acoustic Emission(AE), Conditional Generative Adversarial Nets(CGAN), Rub-Imapct Fault

I. INTRODUCTION

The rotating machinery tends to be more precise and complicated. The gap between dynamic and static parts inside the machinery tends to decrease, which makes dynamic and static parts easily trapped in rub and impact. There are three main research methods for rub impact fault diagnosis of rotating machinery, such as the dynamics analysis of rotor system , vibration analysis and Acoustic Emission (AE) detection[1].The first method focuses on the establishment of dynamic model to study the rub impact mechanism. However it is only suitable for the specific rotor system. It is difficult to obtain satisfactory analysis results for different kinds of rotating machinery[2-3].The vibration analysis method extracts the mechanical vibration characteristics. However, the vibration signals usually show similar characteristics among rub impact faults in rotating machinery. Compared with this problem, the AE signal is much more sensitive to the changes of wear when the dynamic and static parts surfaces broke down.

The fault recognition method, traditional signal processing methods based on Fourier Transform, including Spectrum Analysis, Correlation Analysis, Time Series Analysis, Cepstrum Analysis and Envelope Analysis and so on, have played an important role in fault diagnosis of rotating machinery[4-7].However, prerequisites for these methods are that the input signal must be stationary. By contrast AE signal is a typical non-stationary signal. How to extract comprehensive and effective characteristic to express rub impact fault is a key link of fault identification. Frame features is a method to decompose the complex signal into short-time stabilization.The AE frame features of the rotor can show the details of fault changes from time, frequency and energy intensity views. It is of comprehensive information to diagnosis the fault of rotating machinery.

With the development of Deep Learning, researchers have successively proposed various recognition models based on Deep Learning Networks, such as Convolutional Neural Network (CNN) and Recurrent Neural Networks (RNN)[8-9].Most of these new methods have achieved great success. But deep networks usually rely on large volume samples to learn effective models. However the disassembly of mechanical parts is difficult, data acquisition of dymamic machinery by different equipment and machine tools is not enough. The samples gap remains large for machine learning. In addition, due to serious noises environment the collected rubbing impacted samples are usually contaminated.

Considering available samples not enough, Generative Adversarial Network (GAN) includes generation model to learn the distribution of samples for pseudo-samples construction, discriminant model to evaluate the probability of samples from the training set or real set[10-14]. In the confrontation training between real samples and pseudo-samples, GAN can transform the characteristics of the samples and map into the effective feature subspace for much more robust features, so as to achieve better recognition effect on small data sets. However, GAN usually generates pseudo-samples from random noise, the quality of the quality of generated specific samples is often not guaranteed[15-17].To solve this problem, this paper adopts Conditional Generation

*Aidong Deng is the corresponding author. (e-mail: dnh@seu.edu.cn).

Adversarial Network(CGAN) , which adds useful auxiliary information into each layer of GAN generation model to improve the quality of generated pseudo-samples. The paper proposes a novel rub impact fault recognition method based on CGAN and AE technology so as to achieve better identification effect on small data sets.

II. RUB IMPACT FAULT RECOGNITION BASED ON CGAN

GAN proposed by Goodfellow et al. in 2014 is composed of a generator and a discriminator. The generator captures the potential distribution of real data samples and generates pseudo-samples. A discriminator is a binary classifier that determines whether the input is real data or pseudo-samples. The structure of GAN are shown in Fig. 1. D and G represent discriminator and generator respectively, and input is real data x and random variable z respectively. $G(z)$ is the pseudo-samples generated by G that conforms to the real data distribution as same as possible. The goal of D is to realize data sources. If the sample is real data, it will be marked as 1 for true. If the sample is the production of $G(z)$, it will be marked as 0 for false. Two processes are mutual confrontation and iterative optimization continuously to improve the performance of D and G. Finally it makes the performance of generated pseudo-data $G(z)$ on $D(G(z))$ consistent with the performance of real data x on $D(x)$.
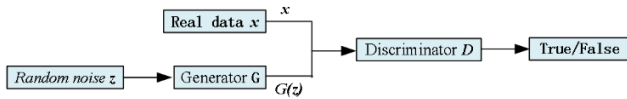
Fig. 1 The Structure of GAN

In the training process , it fixes generator and optimizes discriminator D. The loss function is as following:

$$Obj^D(\theta_D, \theta_G) = -\frac{1}{2}E_{x \sim p_{data}(x)}[\log D(x)] - \frac{1}{2}E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (1)$$

For any non zero real numbers m and n, and $y \in [0,1]$ the expression:

$$-m\log(y) - n\log(1-y) \quad (2)$$

the minimum at $\frac{m}{m+n}$ .Therefore, the target function of generator G is:

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)} \quad (3)$$

According to equation (3), GAN gets the ratio result of the two different probability distributions.

$D(x)$ of x is the probability that x comes from real data, and the goal of G is to get 1. The loss function of generator G is $Obj^G(\theta_G) = -Obj^D(\theta_D; \theta_G)$. Therefore, GAN optimization problem is a minimal and maximization problem.

GAN's objective function can be described as follows:

$$\min_G \max_D \{f(D,G) = E_{x \sim p_{data}(x)}[\log D(x)] + E_{z \sim p_z(z)}[\log(1 - D(G(Z)))]\} \quad (4)$$

The above traditional GAN model can only learn one type of data at once. Consider the practical data sample sets containing multiple classes, it is necessary to learn and generate pseudo sample sets of corresponding classes one by one. To solve this low efficiency problem, the CGAN model structure is proposed in Fig. 2.
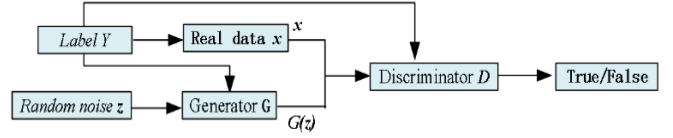

Fig. 2 The Structure of CGAN

It is easy to find out the improvement of model structure. By adding the label information to the generator and discriminator, the GAN model has the ability to generate multiple types of data at once. Besides, at the same time, CGAN model modifies the total loss function, and the new total loss function is shown in equation (5) :

$$\min_G \max_D V(D,G) = \min_G \max_D (E_{x \sim p_{data}(x)}(\ln D(x|Y)) + E_{z \sim p_z(z)}(\ln(1 - D(z|Y)))) \quad (5)$$

Model training processing adopts alternate optimization method: 1) It fixes generator G to optimize discriminator D that the accuracy of discriminating $G(z)$ from real data or pseudo-data distribution achieves maximize. 2)It then fixes discriminator D to optimize generator G that train model G achieves minimize. When training CGAN, the parameters of D are generally updated for k times and the parameters of G are updated for 1 time.

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. AE Data Acquisition

Experimental data from Wind Turbine Drive Train test bed in Southeast University National Engineering Research Center of Turbo-generator Vibration. The test bed is shown in Fig. 3 including drive motor, speed torque sensors, bearing, cylindrical and planetary gearbox and load motor. The motor to adjust the rotor speed, AE sensors installed on the surfaces of bearing shaft. Signal acquisition system is AE acquisition system produced by American Physical Acoustics Corporation. The AE data sampled by 500kHz sample frequency, 51,147 points after an AE impact trigger. Then frame features are obtained by using 512 frame length and 256 frame shift.
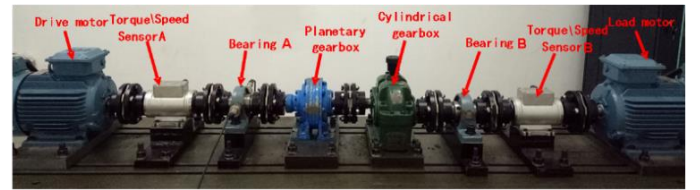

Fig. 3 Test bed setup with AE sensor
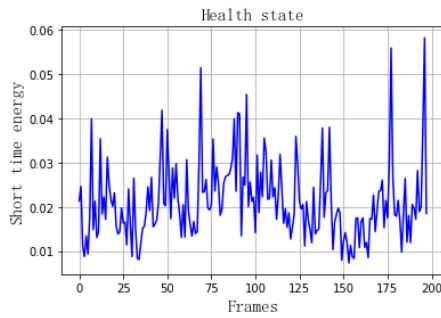
(1) Health state

the

When the bearing runs in health and without rubbing contact, the AE waveform as showing in Fig. 4(a). And also three features based on frame features including spectral skew, short time energy and count above mean are used to recognize rub impact faults in this paper as showing in Fig. 4.It shows that the AE waveform is similar to random noise distribution. In the frame level features normalized short time energy is concentrated at 0.01-0.03. These low-frequency noise spectrum are mainly from the rotor vibration of the mechanical structure.
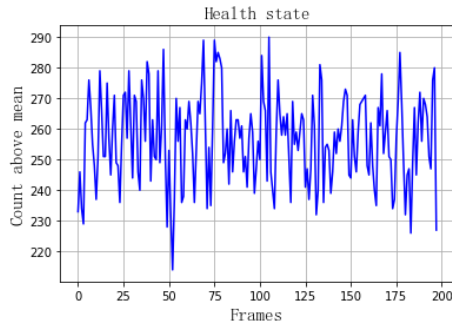
(a) Wavefrom
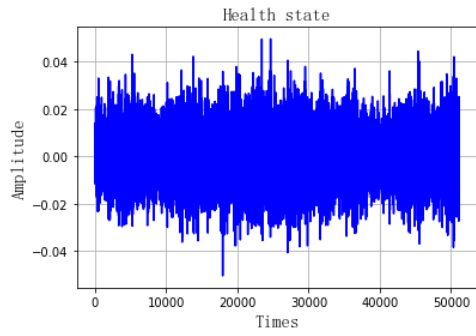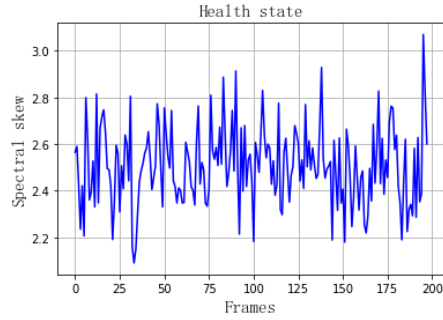
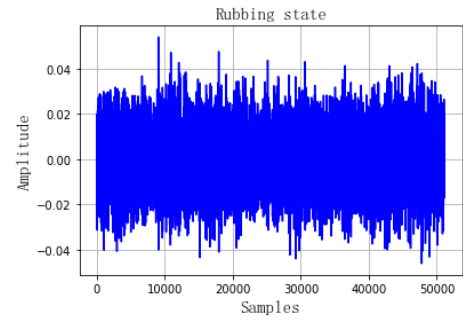(b) Spectral Skew

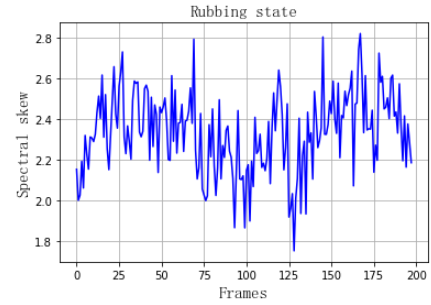(c) Short Time Energy

(d) Count above Mean

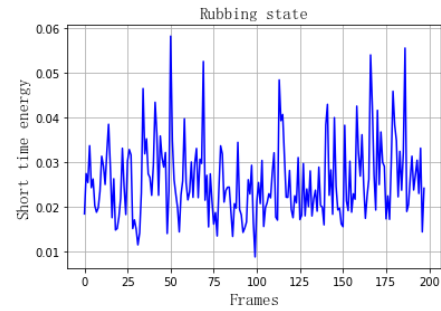Fig.4 AE signal and features at health state

(2) Rubbing state

It changes bearings components that produce full-week rub. Fig.5 shows the AE waveform and its frame features during this rubbing state. From AE waveform , it shows that AE signal fluctuates significantly where the normalized short time energy is concentrated at 0.02-0.04, many AE events occur in the same time. The frequency skew has a noticeable change than in health state.
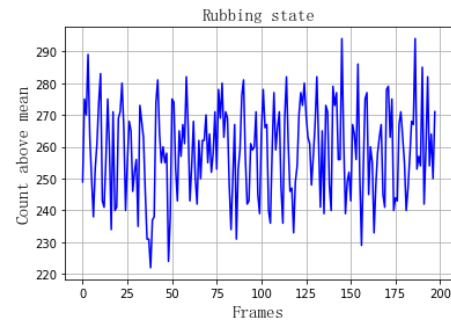

(a) Wavefrom

(b) Spectral Skew
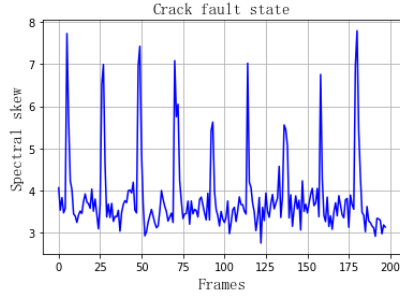

(c) Short Time Energy

(d) Count above Mean

Fig.5 AE signal and features at rubbing state
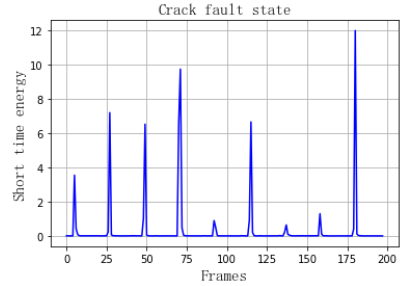
(3) The Crack fault state

It changes bearings components where the inner ring 0.8mm crack fault of bearing was manufactured by spark plug processing technology. As the rotor running, the crack and rotor contacts where local rubbing fault produces AE signal. Fig.6 shows the AE signal and frame features. It can be seen that obvious periodical impact phenomenon in the time space. Also several burst-type AE signals appear in the normalized frequency range of 0-12.
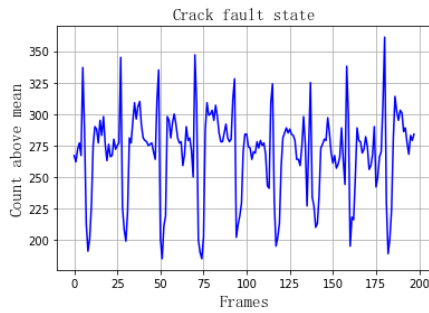
(a) Wavefrom
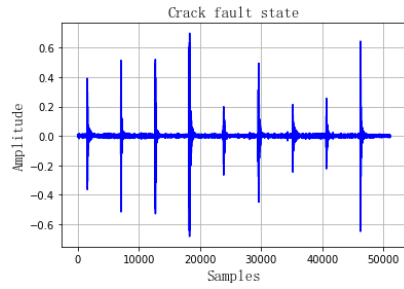
(b) Spectral Skew

(c) Short Time Energy

(d) Count above Mean

Fig.6 AE signal of the Crack fault state

B. CGDN Result Analysis

The samples of the rotor under health state, rubbing state and crack fault state are 700 items in each state. The CGAN network structure is shown in Fig. 7 .
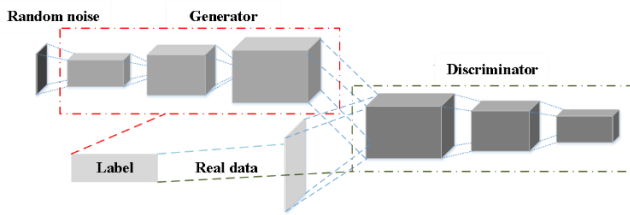

Fig.7 The proposed CGAN network structure

The random noise and the light gray square represents the generator G, which is composed of three layers. Input 100×3 vector after three layers network, generate 594×3 frame features, each layer configuration is as follows:

TABLE1 EACH LAYER OF GENERATOR CONFIGURATION

layers	Input/Output
The first hidden layer	100×3/150×3
The Second hidden layer	150×3/300×3
Output layer	300×3/594×3

The dark gray square represents the discriminator D, which consists of three layers. After three layers the resolution is reduced to size 100.The configuration of each layer is as follows:

TABLE2 EACH LAYER OF DISCRIMINATOR CONFIGURATION

layers	Input/Output
The first layer	594/300
The Second layer	300/150
Output layer	150/100

The parameters are: batch=256, epoch=300, learning rate =0.0001. Adam optimization algorithm was used as the loss function. the The trained CGAN model obtains new sample pairs are formed. New samples are sent to SVM classifier for classification evaluation. The penalty coefficient of SVM is C=1.5, and RBF kernel is adopted.

C. Experiment 1: Performance on Rubbing Fault Recognition Under Ideal Conditions

Under ideal conditions, the performance of fault recognition of rotating machinery with AE signal was analyzed by using CGAN model. The new features obtained through the trained CGAN transformation are mapped into SVM classifier for fault diagnosis. Table 3 is the confusion matrix of using SVM to identify the fault of AE signal characteristics after CGAN conversion. Among them, the recognition result for the crack fault was higher, which was 95.43%.The result of health state recognition was lower, which was 90.54%.The recognition rate of rubbing state was 91.76%.The experimental results show that the pseudo-data generated by the CGAN network provide an effective data supplement for the training of the model. On the other hand, the CGAN network realizes the transformation of the AE features of the AE signal, and adopts the features converted by CGAN. Even the relatively simple linear classifier such as SVM can distinguish the fault state well.

TABLE3 RECOGNITION CONFUSION MATRIX USING CGAN

	Health	Rubbing	Crack
Health	90.54%	4.12%	5.34%
Rubbing	3.01%	91.76%	5.23%
Crack	2.11%	2.46%	95.43%

D. Experiment 2: Performance on Rubbing Fault Recognition in Noise Environment

The experiment adds Gauss white noise into existing AE signal. Table 4 shows the confusion matrix of AE

identification of fault under -20dB noise environment.It can be seen that under the background of noise, the recognition rate of three kinds of faults has decreased. The recognition rate of crack fault is 91.33%, a decrease of 4.1%.In the state of health and rubbing, the recognition rate was 86.14% and 86.49%, decreasing by 4.4% and 5.27%, respectively.

TABLE4 RECOGNITION CONFUSION MATRIX USING CGAN

	Health	Rubbing	Crack
Health	86.14%	5.23%	8.63%
Rubbing	7.52%	86.49%	5.99%
Crack	5.20%	3.47%	91.33%

E. Experiment 3: Performance on Different Identification Methods

The recognition performance of the CGAN model is compared with that of the other two AE fault recognition models commonly used at present. These three common models include: Deep Belief Network(DBN) and GAN.DBN is a three-layer structure and neurons in each layer is 384. Meanwhile, Dropout is used in each layer with coefficient 0.5, and Softmax is used in last layer.The layers structure of GAN are consistent with that of CGAN Network training parameters are set as follows: batch=256, epoch=300, learning rate λ=0.0001. This paper uses Adam optimization algorithm, and cross-entropy loss function.

Table 5 and Table 6 show that the recognition ratio of health state, rubbing state and crack fault state, CGAN model recognition rate increased by 9.5% 1.83% and 1.79% respectively than neural network (DBN), increased by 5.23% 1.43% and 0.76% respectively than GAN model. It further demonstrates that on the similar principle of training and the same number of hidden layer of recognition model, CGAN model has better recognition performance than DBN and GAN model.

TABLE5 RECOGNITION CONFUSION MATRIX USING DBN

	Health	Rubbing	Crack
Health	81.04%	6.76%	1.22%
Rubbing	1.87%	89.93%	0.82%
Crack	3.15%	3.21%	93.64%

TABLE6 RECOGNITION CONFUSION MATRIX USING GAN

	Health	Rubbing	Crack
Health	85.31%	6.52%	8.17%
Rubbing	5.11%	90.33%	4.56%
Crack	2.37%	2.96%	94.67%

IV. CONCLUSION

This paper presents a rub-impact fault diagnosis method using the Conditional Generative Adversarial Nets (CGAN)framework, which solves the problem of the samples gap and improves the quality of generated pseudo samples. It adds classification information into each layer of GAN generation model to improve the quality of the generated pseudo-samples.The data are achieved from the Wind Turbine Train test bed.AE frames features are extracted from various views such as time, frequency and energy intensity under different operation state.Compared with the ideal and noise environments, this model can effectively identify the rub impact fault and have strong robustness.Compared with the other two AE fault recognition models commonly used at present , the experimental results show that the pseudo-data generated by the CGAN network provide an effective data supplement for the training of the model.

REFERENCES

[1] Wei Wang , Weidong Liu , Jing Li ,Wei Peng. A Rub Fault Recognition Method Based on Generative Adversarial Nets[J].Journal of Mechanical Science and Technology, 2020, 34(4):1-10.

[2] Aidong Deng, Hang Tong, Jianeng Tang, Hao Cao, Kang Qin, Xi Yan. Study on Location Algorithms of Beamforming based on MVDR. Applied Mathematics & Information Sciences, 2013, 7(6): 2455-2466.

[3] Deng AD, Zhang XD, Tang JN, Zhao L, Qin K. Localization of Acoustic Emission Source Based on Chaotic Neural Networks.Applied Mathematics & Information Sciences, 2012, 6(3): 713-719.

[4] Feldman M. Hilbert transform in vibration analysis[J]. Mechanical Systems & Signal Processing, 2011, 25(3): 735-802.

[5] Alter O, Brown P O, Botstein D. Singular Value Decomposition for Genome-Wide Expression Data Processing and Modeling[J]. Proceedings of the National Academy of Sciences of the United States of America, 2000, 97(18): 10101-10106.

[6] Park C-S, Choi Y-C, Kim Y-H. Early fault detection in automotive ball bearings using the minimum variance cepstrum[J]. Mechanical systems and signal processing, 2013, 38(2): 534-548.

[7] Joshuva A, Sugumaran V. Wind turbine blade fault diagnosis using vibration signals through decision tree algorithm[J]. Indian Journal of Science and Technology, 2016, 9(48): 1-7.

[8] Mikolov, Tomáš, et al. "Recurrent neural network based language model." Eleventh Annual Conference of the International Speech Communication Association, 2010.

[9] Mohamed A, Dahl G E, Hinton G. Acoustic Modeling Using Deep Belief Networks. IEEE Transactions on Audio Speech & Language Processing, 2011, 20(1): 14-22.

[10] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets[J]. Advances in Neural Information Processing Systems, 2014:2672-2680.

[11] Schlegl T, Seeböck P, Waldstein S M, et al. Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery[M]. Information Processing in Medical Imaging. 2017:146-157.

[12] Kim T, Cha M, Kim H, et al. Learning to discover cross-domain relations with generative adversarial networks[J]. arXiv preprint arXiv:1703.05192, 2017.

[13] Antoniou A, Storkey A, Edwards H. Data Augmentation Generative Adversarial Networks[J]. arXiv preprint arXiv:1711.04340, 2017.

[14] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.

[15] Salimans T, Goodfellow I, Zaremba W, et al. Improved techniques for training gans[C]//Advances in Neural Information Processing Systems. 2016: 2234-2242.

[16] Springenberg J T. Unsupervised and Semi-supervised Learning with Categorical Generative Adversarial Networks[J]. Computer Science, 2015.

[17] Grall A , Hamidinekoo A , Malcolm P , et al. Using a Conditional Generative Adversarial Network (cGAN) for Prostate Segmentation[M]// Medical Image Understanding and Analysis. 2020.

Analysis of Real-time Noise Signal Characteristics of Power Transformer Based on All-phase Fast Fourier Transform

Xiaojuan Zhao*
Grid Technology Center
State Grid Chongqing Electric Power Research Institute
Chongqing, China
1012254736@qq.com

Yaqi Zhou
Planning and Operations Department
State Grid Hubei Transmission and Transformation Engineering CO.LTD
Wuhan, China
zhouyaqi-kathy@163.com

Xue Wang
Postal Savings Department
32620 Army
Xining, China
1075678609@qq.com

Ruilin Xu
Development Safety Supervision Department
State Grid Chongqing Electric Power Research Institute
Chongqing, China
xrl86@163.com

Shihai Han
Energy Internet Technology Center
State Grid Chongqing Electric Power Research Institute
Chongqing, China
han2008@qq.com

Xin Chen
Planning and Operations Department
State Grid Hubei Transmission and Transformation Engineering CO.LTD
Wuhan, China
943409630@qq.com

Abstract—**Analysis of the time and frequency domain characteristics of the real-time noise signal of the transformer is beneficial to the research of active noise control technology in substations. When performing spectrum analysis on transformer noise, the All-phase fast Fourier transform (FFT) spectrum analysis method has more advantages than the traditional analysis method. This method is used to process the real-time noise signal of a transformer. As can be seen from its spectrum chart, the transformer noise is based on 100Hz as the fundamental frequency. The noise energy reaches the peak at 100Hz and 200Hz. In addition, the A sound level of the transformer noise signal can be calculated through the Paseval relationship. The A sound level changes smoothly with time and has a small fluctuation amplitude. Although the transformer is a large and complex noise source, the primary sound field generated by it is stable and conducive to implementation of active noise control system.**

Keywords—Transformer, noise signal, All-phase FFT spectrum analysis, A-weighted sound level

I. INTRODUCTION

The analysis and processing of audio signals have been widely used in the fields of noise suppression, non-destructive testing of materials and recognition of speech signals. At present, a variety of audio signal analyzers are used in actual projects at home and abroad. Although these instruments can measure multiple performance indicators of audio signals, they are expensive and not targeted, which is not conducive to the analysis of audio signal characteristics. In practical application, studying the time domain and frequency domain characteristics of transformer noise signals is of great significance to substation noise control [1,2]. The methods for discrete spectrum analysis of signals include traditional FFT method, Welch method, and multi-pyramid method. However, when these spectrum analysis methods process long signal sequences, the data need to be truncated, causing serious spectral leakage. For this reason, the All-phase fast Fourier transform (FFT) spectrum analysis method came into being. This analysis method can not only effectively suppress side spectrum leakage, but also have "phase invariance" [3]. At present, All-phase FFT spectrum analysis and its correction

method have been widely used in the precise adaptive notch in noisy environments, continuous wave radar speed measurement, harmonic analysis of power systems and the design of spectrum analyzers [4]. In view of the above reasons, this paper directly collects the real-time noise signal of the transformer and performs All-phase FFT spectrum analysis on it to obtain the spectrogram of the noise signal. In addition, the paper calculates the A-weighted sound level of the real-time noise signal through the Parseval relationship, measures the subjective perception of noise in the human ear and studies the acoustic characteristics of transformer noise from two aspects of time domain and frequency domain.

II. FREQUENCY DOMAIN CHARACTERISTICS OF TRANSFORMER NOISE SIGNAL

In digital signal processing, the discrete Fourier transform (DFT) can be used to obtain the frequency domain characteristics of the signal and its calculation formula is shown in (1).

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} \qquad (k = 0,1,\cdots,N-1) \qquad (1)$$

$X(n)$ is the input discrete digital signal sequence, $X(k)$ is the relative amplitude at the kth discrete frequency point. Although DFT can obtain the frequency domain characteristics of the signal, its large amount of calculation and long time consumption are not conducive to real-time signal processing [5]. In practical applications, the improved algorithm is generally used to analyze the signal spectrum.

A. Traditional Spectrum Analysis Methods

Traditional FFT method is used to describe the distribution of signal power over frequency. For the signal sequence $x_L(n)$ of length L, the spectral characteristics of N sampling points are analyzed to obtain the spectral amplitude value [6].

$$\hat{P}_x(f_k) = \frac{|X_L(f_k)|^2}{f_s L} \qquad (2)$$

$$f_k = \frac{kf_s}{N} \qquad (k = 0,1,\cdots,N-1) \qquad (3)$$

It can be seen from (2) and (3) that as the number of sampling points increases, the resolution is higher, but the limitations of the sequence length lead to spectral leakage in this analysis method. In addition, the expected value and variance value can be calculated separately.

$$E\left[\hat{P}_x(f_k)\right] = \frac{1}{f_s L}\int_{-f_s/2}^{f_s/2} P_x(q)|W(f_k - q)|^2\, dq \tag{4}$$

$$D\left[\hat{P}_x(f_k)\right] \approx P_x^2(f_k)\left\{1 + \left[\frac{\sin(2\pi L f_k / f_s)}{L\sin(2\pi f_k / f_s)}\right]^2\right\} \tag{5}$$

$|W(f_k - q)|^2$ is the convolution of two rectangular windows. Equation (4) shows that the traditional FFT method is progressive and unbiased. As the data length increases, the resulting spectrum becomes more and more accurate. Due to its large variance, the resulting spectral characteristic curve fluctuates sharply. It can be seen from (5) that even if the data length is increased, the variance cannot be greatly reduced.

In order to improve the variance performance of the traditional FFT method, the data sequence is first divided into overlapping data segments and the periodic graph spectral estimation of each segment of data is windowed and averaged, which is the Welch spectrum analysis method [7]. The expected value of the Welch method is calculated as in (6).

$$E\left[\hat{P}_w(f_k)\right] = \frac{1}{f_s L_s U}\int_{-f_s/2}^{f_s/2} P_x(q)\left|W(f_k - q)\right|^2 dq \tag{6}$$

L_s is the length of the segmented data. $U = \frac{1}{L}\sum_{n=0}^{L-1}|w(n)|^2$ is the window normalization constant. For a certain length of data, due to $L > L_s$, the deviation estimated by the Welch method will be greater than the traditional FFT method, and its variance is related to the segment length and window function, which is difficult to quantify.

Thompson's Multi-Vertex Method (MTM) provides a time-bandwidth parameter M, which is used to balance the estimated variance and resolution. This parameter is related to time, bandwidth, and multi-vertebral number. As M increases, the variance decreases. However, the bandwidth of each multi-vertebral body is proportional to M. Increasing M will cause greater spectral leakage. For each set of data, there is always a parameter M to obtain the best performance between the estimated deviation and the variance.

B. All-phase FFT Spectrum Analysis

Compared with the above-mentioned various spectrum analysis methods, the All-phase FFT (apFFT) spectrum analysis method has the advantages of unchanged phase, suppressed side lobe leakage and reduced variance [8-10]. Before performing apFFT spectrum analysis, the data should be pre-processed [3, 8]. The specific implementation steps are as follows.

1) A data vector of length **(2N-1)** is transformed into **N** vectors of length N. This vector group should contain all the segmentation of sample points $x(n)$ and the mapping relationship is shown in (7).

$$\boldsymbol{x}_i = \left[x(n+i-N+1),\cdots,x(n),\cdots,x(n+i)\right]^T \tag{7}$$

Where, $i = 0,1,\cdots,N-1$.

2) The sequence $\boldsymbol{x}_i$ is weighted with the front window sequence $\boldsymbol{f} = [f(0),...,f(i),...,f(N-1)]$ and the sequence after the window is extended. $x(n)$ term is in the first position. The back window sequence $\boldsymbol{b} = [b(0),...,b(i),...,b(N-1)]^T$ is used to weight the new sequence to obtain the vector group as shown in (8).

$$\boldsymbol{y}_i = \left[x(n)\cdot f_{N-1-i},\cdots,x(n+i)\cdot f_{N-1},x(n+i-N+1)\cdot f_0,\cdots,x(n-1)\cdot f_{N-2-i}\right]^T\cdot b_i \tag{8}$$

3) The corresponding elements of the above vector groups are summed and averaged to form a All-phase input sequence to realize data preprocessing.

The process of All-phase data preprocessing can be equivalent to a long $(2N-1)$ window sequence $\boldsymbol{w}_c$ to weight the data vectors and then shift and superimpose to obtain All-phase input sequence [10]. The FFT of the input sequence is the output of apFFT. In order to ensure that the amplitude of the input sequence and the output sequence do not deviate, the window sequence $\boldsymbol{w}_c$ should be normalized. If the window function $\boldsymbol{w}$ is formed by convolution of the front window sequence $\boldsymbol{f}$ and the back window sequence $\boldsymbol{b}$, the calculation formula is shown in (9).

$$w(n) = f(n)*b(-n) = \begin{cases} \sum_{k=0}^{N-1-n} b_k f_{k+n} & n \in [0, N-1] \\ \sum_{k=-n}^{N-1} b_k f_{k+n} & n \in [-N+1,-1] \\ 0 & others \end{cases} \tag{9}$$

Then the normalized convolution window $\boldsymbol{w}_c$ can be calculated by (10).

$$w_c(n) = w(n) / \sum_{n=-N+1}^{N-1} w(n) \tag{10}$$

Preprocessing the signal sequence with this overlapping method can improve the side lobe leakage caused by the traditional spectrum analysis method. The single frequency complex exponential signal $x(n) = e^{j(n\omega_0 + \varphi_0)}$ is taken as an example. $\omega_0 = 2\beta\pi / N$. Then the output of the windowless traditional FFT and windowless apFFT of the sequence $\{x(n)\}$ are shown in (11) and (12) respectively.

$$X(k) = \frac{1}{N}\cdot\frac{\sin[\pi(\beta-k)]}{\sin[\pi(\beta-k)/N]}\cdot e^{j[\varphi_0 + \frac{N-1}{N}(\beta-k)\pi]} \tag{11}$$

$$Y(k) = \frac{e^{j\varphi_0}}{N^2}\cdot\frac{\sin^2[\pi(\beta-k)]}{\sin^2[\pi(\beta-k)/N]} \tag{12}$$

Equations (11) and (12) indicate that the All-phase FFT spectral amplitude of the unit complex exponential signal is the square of the traditional FFT spectral amplitude, which means that the ratio of the side spectral line relative to the main spectral line is also attenuated according to the square relationship, so apFFT can better suppress sidelobe leakage [4]. In addition, the phase value of each spectral line of the All-phase FFT is φ_0, which is the theoretical phase value of the center sample point $x(0)$. Because the traditional spectrum analysis method has the disadvantages of large variance, serious side lobe leakage and inability to extract the phase value, the All-phase FFT spectrum analysis method has
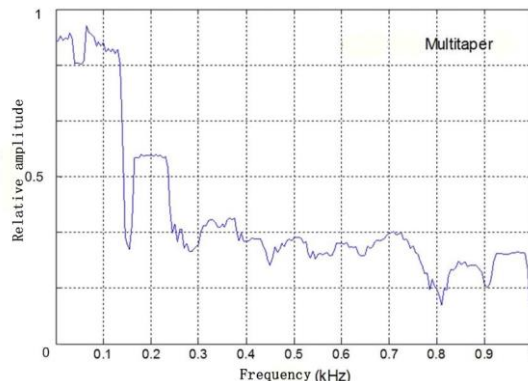
advantages in analyzing the frequency domain characteristics of the transformer noise signal.
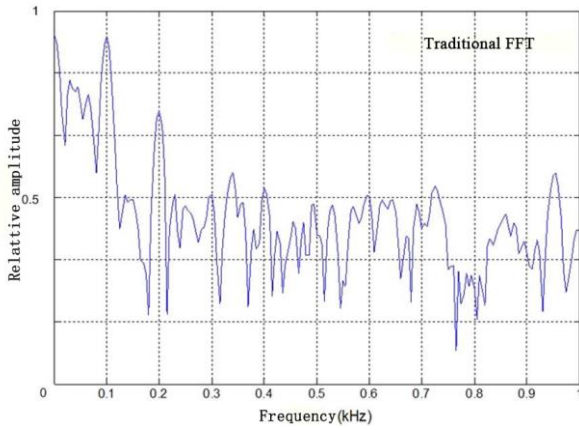
(a)The frequency characteristic curves by traditional FFT
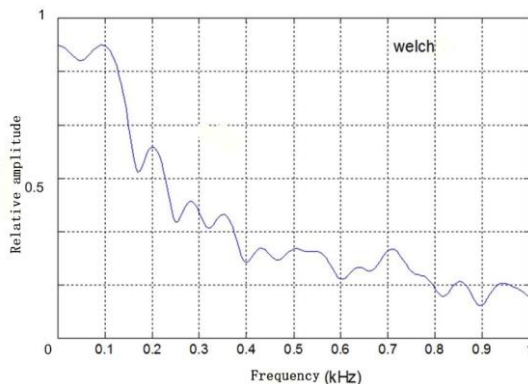
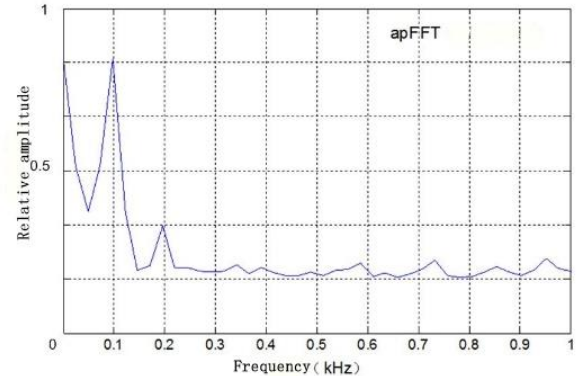(b)The frequency characteristic curves by welch

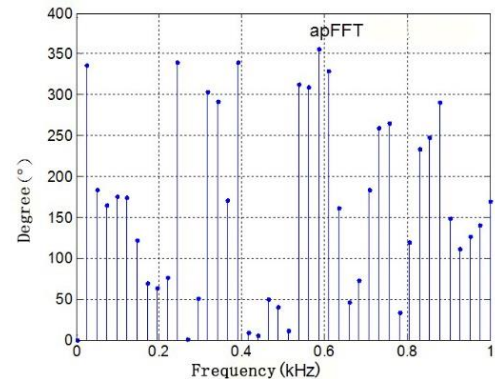(c)The frequency characteristic curves by multitaper

C. Spectrum Analysis of Transformer Noise Signal

The traditional FFT method, the welch method, the multi-pyramid method and the apFFT spectrum analysis method are used to perform spectrum analysis on the acquired transformer noise signal. The spectrum diagram is obtained as shown in Fig. 1.

Combined with various spectrum analysis methods, it can be seen that the transformer noise signal takes 100 Hz as the fundamental frequency and reaches the energy peak at the frequency points of 100 Hz and 200 Hz. Among them, the curve obtained by the traditional FFT method is not smooth and the variance is large. In contrast, the welch method can

(d)The frequency characteristic curves by apFFT

(e)The degree by apFFT

Fig. 1.   The frequency characteristic curves of transformer noise

improve the defect of the large variance of the traditional FFT method, but the resolution is significantly reduced. Although the multi-pyramid method balances the relationship between the variance and the resolution, its calculation amount is large, which is not conducive to the processing of real-time noise signals. The All-phase FFT spectrum analysis method has a higher main side lobe ratio, which can better suppress side lobe leakage and reduce variance, which is beneficial to the extraction of transformer noise spectrum characteristics. In addition, apFFT has "phase invariance". The phase of the characteristic frequency point can be extracted and used in the noise reduction system to control the placement and parameter adjustment of the secondary sound source.

III.   A-WEIGHTED SOUND LEVEL SYNTHESIS

The A-weighted sound level of the real-time noise signal of the synthesized transformer can be used to measure the sensitivity of the human ear to noise, intuitively reflect the effect of noise control and have a major role in the control of transformer noise. In addition, observing the change of the noise A-weighted sound level of the transformer noise signal with time is conducive to the study of its time-domain characteristics and provides a theoretical basis for noise control.

A. A-weighted Network Filtering

In noise measurement and evaluation, the A-weighted network is generally used to correct different frequency components in the signal. According to the frequency weighting characteristics of A-weighted sound level, the filter

coefficients of the weighted network at different frequencies are fitted, as shown in (13) [11].

$$\alpha_A = \frac{(3.5 \times 10^{16}) f^8}{[(20.6^2 + f^2) \cdot (107.63^2 + f^2) \cdot (737.86^2 + f^2) \cdot (121942.2^2 + f^2)]^2}$$

(13)

The filter coefficient reaches a peak value of 1.34 at a frequency around 2.5kHz, and the weighting network has the effect of enhancing the signal between 1kHz and 6.1kHz. For other frequency bands, the signal has different degrees of attenuation.

B. Calculate A-weighted Sound Level

The energy of the transformer noise can be expressed by the square sum of the amplitude of the time domain signal, but after the All-phase FFT, the energy of the signal in the time domain needs to be calculated by inverse transformation.

Paceval's theorem states that the energy of a signal is conserved in the time and frequency domains [12]. Therefore, it is more convenient to choose to calculate the energy of the signal in the frequency domain, as shown in (14).

$$E = \sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N} \sum_{i=0}^{N-1} |X(i)|^2$$

(14)

$X(i)$ is the discrete Fourier transform of the sequence $x(n)$. According to (11) and (12), the relationship between apFFT and DFT output is shown in (15).

$$|Y(i)| = |X(i)|^2$$

(15)

Combining the operation result of the filter coefficient α_A of the A-weighted network and the All-phase FFT, the output of the A-weighted All-phase FFT is shown in (16).

$$Y_A(i) = \alpha_A(f_i) \cdot Y(i) \qquad i = 0, 1, \cdots, N-1 \quad (16)$$

Where: $f_i = i \cdot \Delta f = i \cdot \dfrac{f_s}{N}$.

Then the energy of the noise after passing through the A-weighted network is shown in (17).

$$E_A = \frac{1}{N} \sum_{i=0}^{N-1} (\alpha_A \cdot |X(i)|)^2 = \frac{1}{N} \sum_{i=0}^{N-1} |Y_A(i)| \cdot \alpha_A$$

(17)

The average energy per unit time is shown in (18).

$$e \approx \frac{1}{N \Delta t} \sum_{i=0}^{N-1} |Y_A(i)| \cdot \alpha_A$$

(18)

In order to determine the real-time A-weighted sound level of the transformer noise signal, there must be a reference signal. Assume that its average energy is e_0, then

$$dBA = 10(\lg e - \lg e_0)$$

(19)

$10 \lg e_0$ is a constant, so the signal A-weighted sound level can also be expressed as in (20).

$$dBA = 10 \lg e + C$$

(20)

The calibration constant C can be obtained by experiment, the paper uses known sample signal to adjust the constant C.

IV. APPLICATIONS

The real-time noise of a transformer is taken as an example. Noise signals are collected to analyze their frequency and time domain characteristics.

A. Analysis of Real-time Noise Signal of Transformer

When using All-phase FFT to perform spectrum analysis on the real-time noise signal of the transformer, the number of samples N_a should be determined by the sampling rate f_s and sampling time interval Δt, as shown in (21).

$$N_a = (2N - 1) = f_s \times \Delta t$$

(21)

Since the highest frequency of human audible sound is 20kHz and the Nyquist frequency supported by the sound card is selected to be 20.5kHz, the actual sampling rate f_s is 41kHz. If the sampling time interval Δt is 0.15s, it can be calculated that N_a is 6150, but when apFFT is used to perform spectral analysis on N_a data, the N data should be subjected to FFT spectral analysis, then N should be a power index of 2. Sample number N is power exponent 4096 slightly larger than $(N_a + 1)/2$. Then the actual sampling number N_a is 8191. Since the sampling rate f_s is 41 kHz, the actual time interval Δt is about 200 millisecond.

The real-time noise analysis of a transformer is performed from the three aspects of apFFT spectrum analysis, real-time audio signal and A-weighted sound level synthesis. The flow chart is shown in Fig. 2.

The real-time noise signal of a transformer is collected and converted into an electrical signal and displayed. Electrical signals are All-phase data preprocessed and performed by apFFT spectrum analysis. Spectrogram of noisy signal is shown. In addition, the noise energy is calculated according to the results of the calibration constant and apFFT, and then the transformer noise A-weighted sound level is synthesized and the decibel number is displayed. The display interface is updated in real time and the analysis result at a certain time is shown in Fig. 3.
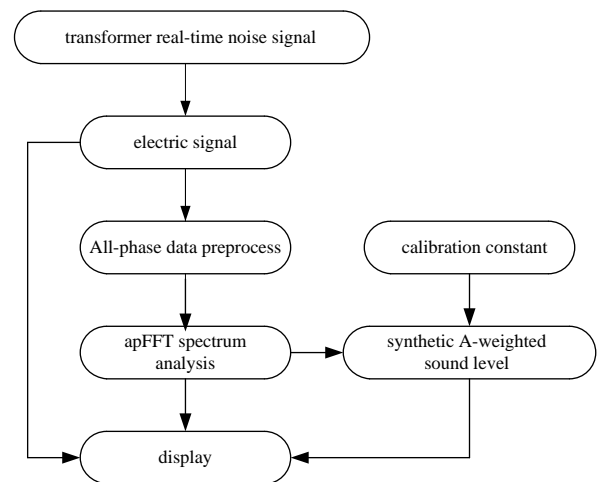

Fig. 2. The flow diagram of real-time transformer noise signals analysis

The spectrum characteristic curve of the transformer noise under test has little change with time. As can be seen from Fig. 3, the energy of the noise signal is mainly concentrated in the

low frequency band. When the frequency is greater than 1 kHz, the noise energy decay faster. In addition, the collected real-time noise signal changes smoothly over time with little fluctuation. It is shown that the noise A-weighted sound level at this moment is 58.0dB.
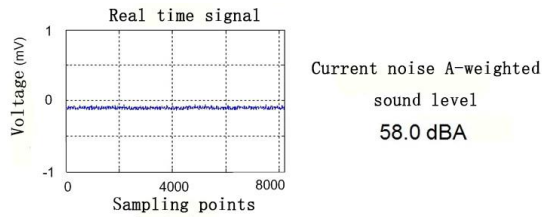
(a) The real-time transformer noise signals
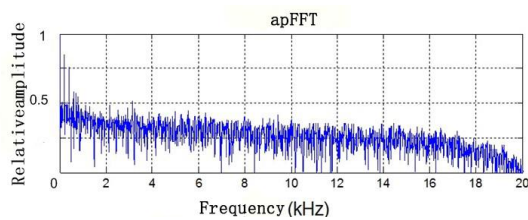
(b) The spectrum characteristic curve of the transformer noise

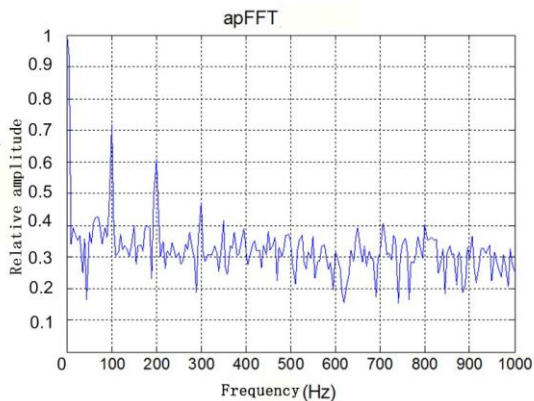Fig. 3. An analysis of real-time transformer noise signals

Fig. 4. The frequency characteristic curves of real-time transformer noise signals

B. *Characteristic Analysis of Noise Signal*

The spectrum analysis range is limited to within 1 kHz. The apFFT spectrum analysis of the transformer noise signal is performed respectively. The spectrum diagram is obtained as shown in Fig. 4.

It can be obtained from the spectrum characteristic curve in Fig. 4 that the noise of the transformer under test takes 100 Hz as the fundamental frequency, and the energy peak is larger at the frequency points of 100 Hz and 200 Hz. It can be seen

that the noise signal changes smoothly with time by observing the real-time noise signal of the transformer. The change of the A-weighted sound level is less than 1dB within 1min. Although the transformer is a large and complex noise source, the primary sound field it generates is relatively stable, which is beneficial to the implementation of active noise control systems.

V. CONCLUSION

1) The performance difference between the traditional spectrum analysis method and the All-phase FFT spectrum analysis method is compared. ApFFT has a smaller variance and truly reflects the initial phase of the original signal.

2) The apFFT spectrum analysis of the real-time noise signal of the transformer can be obtained. The energy of the noise signal is mainly concentrated in the middle and low frequency bands and the fundamental frequency is 100Hz. The energy peak is larger at the frequency points of 100Hz and 200Hz.

3) It can be seen that the noise signal is stable and the fluctuation amplitude is small through observing the real-time noise signal of the transformer and the change of it's A-weighted sound level with time.

REFERENCES

[1] Teoh C, Soh K, and Zhou R. "Active noise control of transformer noise," [Proceedings of the 1998 2nd International Conference on Energy Management and Power Delivery, p. 747-753, 1998].

[2] Wu Xiaoxuan, Yang Yaping, and Liu Jian, "Noise test and characteristic analysis of 500kv substation," East China Electric Power, 2010, 38(6), pp. 887-889.

[3] Wang Zhaohua, Huang Xiangdong, Full phase spectrum analysis and filtering technology. 2nd ed, Beijing: Electronic Industry Press, 2009.

[4] Huang Xiangdong, All-phase digital signal processing. Tianjin: Tianjin University, 2006.

[5] Lv Wentao, Tian Zhongshan. "Real-time spectrum analyzer based on all phase FFT spectrum analysis". [Proceedings of the 2013 4th International Conference on Digital Manufacturing and Automation, p. 966-969, 2013].

[6] Xue Xiaopeng, Zhao Jun, and Zhang Xinyi, "Spectral analysis of the two-dimensional hydrodynamic characteristics of the Sanchahe sluice and the pulsating pressure on the gate," Hydropower Energy Science, 2014, 32(1), pp. 180-183.

[7] Sun Jianping, Dai Yongfeng, "Noise analysis of 21F unit of Three Gorges Power Plant during high load operation," Hydropower Energy Science, 2011, 29(5), pp. 117-119.

[8] Sun Shanlin, Research on key technologies of all-phase OFDM system. Tianjin: Tianjin University, 2009.

[9] Tan Siwei, Ren Zhiliang, "Improvement of phase difference correcting spectrum method based on all-phase FFT," Systems Engineering and Electronics, 2013, 35(1), pp. 34-39.

[10] Xu Nini, All-phase FIR filter bank. Tianjin: Tianjin University, 2005.

[11] He Lingsong, "Formulated expression of frequency weighting network," Journal of Instrumentation, 1996, 17(5), pp. 540-542.

[12] Ma Dayou, Shen Hao. Acoustic Handbook. Beijing: Science Press, 1983.