Lecture Notes in Electrical Engineering 334

Wego Wang *Editor*

Proceedings of the Second International Conference on Mechatronics and Automatic Control

Springer

Lecture Notes in Electrical Engineering

Volume 334

Lecture Notes in Electrical Engineering (LNEE) is a book series which reports the latest research and developments in electrical engineering, namely:

- Communication, networks, and information theory
- Computer engineering
- Signal, image, speech and information processing
- Circuits and systems
- Bioengineering

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer's high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer's other Lecture Notes series, LNEE will be distributed through Springer's print and electronic publishing channels.

To submit a proposal or request further information, please contact the appropriate Springer Publishing Editors:

Asia:

China, *Lu Yang, Editor* (lu.yang@springer.com)
(Electrical Engineering)

Japan, *Yuko Sumino, Editor* (yuko.sumino@springer.com)
(Electrical Engineering)

Southeast Asia, *Ramesh Premnath, Editor* (ramesh.premnath@springer.com)
(Electrical Engineering)

South Asia, *Aninda Bose, Editor* (aninda.bose@springer.com)
(Electrical Engineering)

Europe:

Thomas Ditzinger, Editor (thomas.ditzinger@springer.com)
(Robotics & Control, Computational Intelligence & Complexity, Mechanics and Computational Methods)

Christoph Baumann, Editor (christoph.baumann@springer.com)
(Heat and Mass Transfer, Signal Processing and Telecommunications, and Solid and Fluid Mechanics, and Engineering Materials)

North America:

Michael Luby, Editor (michael.luby@springer.com)
(Mechanics; Materials)

**** Now indexed by ISI Proceedings, EI-Compendex, SCOPUS, EI Compendex, MetaPress, Springerlink ****

More information about this series at http://www.springer.com/series/7818

Wego Wang
Editor

Proceedings of the Second International Conference on Mechatronics and Automatic Control

Springer

Editor
Wego Wang
University of Massachusetts Lowell
Wellesley
MA
USA

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-3-319-13706-3 ISBN 978-3-319-13707-0 (eBook)
DOI 10.1007/978-3-319-13707-0

Library of Congress Control Number: 2015943319

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Part VIII Control and Optimization

Contributors

Weiwei An Nanjing Institute of technology, Nanjing, China

Xueli An China Institute of Water Resources and Hydropower Research, Beijing, China

ShunXian Bai School of Electronics and Information Engineering, North National University, Yinchuan, China

Xiaojuan Bai 95997 Troops of Chinese PLA, Beijing, China

Le Bao Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University), Ministry of Education, Jinan, China

Haihong Bian School of Electronic Power Engineering, Nanjing Institute of Technology, Nanjing, China

Pham Hai Binh Hanoi Technology University, Hanoi, Vietnam

Dehua Cai JiangMen Power Bureau, Jiangmen, China

Xi Cai Key Laboratory of Dynamics and Control of Flight Vehicle, Ministry of Education, School of Aerospace Engineering, Beijing Institute of Technology, Beijing, China

Jianfeng Cao National Key Laboratory of Science and Technology on Aerospace Flight Dynamics, Beijing Aerospace Control Center, Beijing, China

Zhengdi Cao School of Electrical Engineering, Shenyang University of Technology, Shenyang, China

Ye Chai State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Zhejiang, China

Chen Chen School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, China

Hanning Chen Shenyang Institute of Automation Chinese Academy of Sciences, Shenyang, China

Honggang Chen Gansu Electric Power Corporation, Lanzhou, China

Jiahong Chen State Grid Electric Power Research Institute, Wuhan, China

Jie Chen School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing, China

Kangting Chen School of Economics and Management, North China Electric Power University, Beijing, China

Ke Chen State Key Laboratory of ISN, Shenzhen CU-Xidian joint center, Xidian University, Xi'an, China

Liang Chen Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, Wuhan, China

Lue Chen National Key Laboratory of Science and Technology on Aerospace Flight Dynamics, Beijing Aerospace Control Center, Beijing, China

Tao Chen National ASIC System Engineering Research Center, Southeast University, Nanjing, China

Weiyu Chen Electric Power College, South China University of Technology, Guangzhou, China

Xiaolong Chen The School of Electro-Mechanical Engineering, Xidian University, Xi'an, China

Xiaoyue Chen School of Electrical Engineering, Wuhan University, Wuhan, China

Yihuang Chen School of Electrical Engineering, Wuhan University, Wuhan, China

Ying Chen Institute of Optoelectronic Technology, Beijing University of Aeronautics and Astronautics, Beijing, China

Yuqiang Chen Department of Computer Engineering, Dongguan Polytechnic, Guangdong, China

Dong Cheng College of Electrical and Information Engineering, Hunan University, Changsha, China

Hong Cheng Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, Wuhan, China

Lefeng Cheng Electric Power College, South China University of Technology, Guangzhou, China

Beihua Cong Tobacco Guangdong Industrial Co., Ltd., Guangzhou, China

Wei Cong Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University) Ministry of Education, Jinan, China

Wei Cong Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University), Ministry of Education, Jinan, China

Balla Moussa Coulibaly School of Electrical and Electronic, Engineering North China Electric Power University, Beijing, China

Jianhua Cui Shanghai Regeon Electric Co., Ltd., Shanghai, China

Manyuan Dai Electronic Science and Engineering Department, National University of Defense Technology, Changsha, China

Yingjie Dai Mechanical Engineering College, Shijiazhuang, China

Guiping Deng Metering Center, State Grid Hubei Electric Power Research Institute, Wuhan, China

Hailong Deng Logistics College of PLA, Beijing, China

Shijie Deng Mechanical Engineering College, Shijiazhuang, China

Yaping Deng School of Automation, Huazhong University of Science and Technology, Wuhan, China

Zhao Deng Department of Aeronautics, School of Physics and Mechanical & Electrical Engineering, Xiamen University, Xiamen, China

Haiting Diao College of Information Science and Academic, Shandong Agricultural University, Tai'an, China

Huijuan Ding Department of Electronic and Information Engineering, Inner Mongolia University, Hohhot, China

Li Ding Metering Center, State Grid Hubei Electric Power Research Institute, Wuhan, China

Chao Dong College of Information Science and Academic, Shandong Agricultural University, Tai'an, China

Lei Dong School of Automation, Beijing Institute of Technology, Beijing, China

Shuai Dong Automation Research and Design Institute of Metallurgical Industry, Beijing, China

Tianyu Dong School of Electrical Engineering, Shenyang University of Technology, Shenyang, China

Yangxia Dong School of Automation, Northwestern Polytechnical University, Xi'an, China

Huijie Du State Grid Shanxi Electric Power Research Institute, Taiyuan, China

Y ongchang Du State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing, China

Rong Fan School of Instrument Science and Opto-electronics Engineering, Hefei University of Technology, Hefei, China

Chonghua Fang Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Centre, Wuhan, China

Fanxiu Fang Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University) Ministry of Education, Jinan, China

Weidong Fang Shanghai Institute of Microsystem and Information Technology, CAS, Shanghai, China

Yuan Fang College of Mechanical and Electrical, Changchun University of Science and Technology University, Changchun, China

Qinhua Feng Qingyuan Power Supply Company, Qingyuan, China

Shufang Feng Xinxiang Xinovo Machinery Co., Ltd, Xinxiang, China

Xin Fu State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Zhejiang, China

Xisong Fu State Key Laboratory of ISN, Shenzhen CU-Xidian joint center, Xidian University, Xi'an, China

Ge Gao Electronic Science and Engineering Department, National University of Defense Technology, Changsha, China

Hengxuan Gao BeiHang University, Beijing, China

Pu Gao School of Vehicle and Energy Engineering, Yanshan University, Hebei, China

Rong Gao State Key Laboratory of ISN, Shenzhen CU-Xidian joint center, Xidian University, Xi'an, China

Weiwei Gao Shanghai IoT Co., LTD., Shanghai Research Center for Wireless Communications, Shanghai, China

Luo Ge School of Electrical Engineering, Wuhan University, Wuhan, China

Suiyan Geng School of Electrical and Electronic, Engineering North China Electric Power University, Beijing, China

Cheng Gong School of Mathematical Science, Heilongjiang University, Harbin, China

Fang Gu School of Physics and Optoelectronic Engineering, Nanjing University of Information Science and Technology, Nanjing, China

Huaxi Gu State Key Laboratory of ISN, Shenzhen CU-Xidian joint center, Xidian University, Xi'an, China

Ai Guo Sichuan Institute of Solid-State Circuits, Chongqing, China

Rui Guo Laser Research Institute, Shandong Academy of Science, Jinan, China

Siyu Guo College of Electrical and Information Engineering, Hunan University, Changsha, China

Weilong Guo State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Zhejiang, China

Yue Guo Metering Center, State Grid Hubei Electric Power Research Institute, Wuhan, China

Le Thi Thanh Ha Hanoi Technology University, Hanoi, Vietnam

Jianfeng Han School of Information Engineering, Inner Mongolia University of Technology, Huhhot, China

Songtao Han National Key Laboratory of Science and Technology on Aerospace Flight Dynamics, Beijing Aerospace Control Center, Beijing, China

Sulan Han North China University of Water Resources and Electric Power, Zhengzhou, China

Xu Han Naval Aeronautical and Astronautical University, Yantai, China

Wei Hao College of Electrical Engineering & New Energy, Three Gorges University, Yichang, China

Xiaohong Hao College of Computer and Communcation, Lanzhou University of Technology, Lan zhou, China

Yingguang Hao Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China

Jun He State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Zhejiang, China

Long He The Aeronautic Maintenance Engineering Department, Chengdu Aeronautic Vocational and Technical College, Chengdu, China

Ruofei He No. 365 Research Institute, Northwestern Polytechnical University, Xi'an, China

Tian p eng He BeiHang University, Beijing, China

Xiaobo He School of Control and Computer Engineering, North China Electric Power University, Beijing, China

Zhiqiang He Hunan Electric Power Test & Research Institute, Changsha, China

Rongjing Hong School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing, China

Shuai Hou School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China

Guozhao Hu College of Automation, Wuhan University of Technology, Wuhan, China

Jingfang Hu College of Communication Engineering, Hangzhou Dianzi University, Hangzhou, China

Kunyuan Hu Shenyang Institute of Automation Chinese Academy of Sciences, Shenyang, China

Liang Hu State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Zhejiang, China

Ming Hu School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan, China

Weitao Hu State Grid Hebei Electric Power Company, Shijiazhuang, China

Xiao Hu 95997 Troops of Chinese PLA, Beijing, China

Xinrong Hu School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan, China

Yanpeng Hu BeiHang University, Beijing, China

Hao Huang North China Electric Power University, Beijing, China

Minfang Huang School of Economics and Management, North China Electric Power University, Beijing, China

Ping Huang Chengdu Neusoft university, Chengdu, China

Jianfeng Jia School of Control and Computer Engineering, North China Electric Power University, Beijing, China

Chen Jiang College of Computer and Communcation, Lanzhou University of Technology, Lan zhou, China

Minghua Jiang School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan, China

Wuxue Jiang Department of Computer Engineering, Dongguan Polytechnic, Guangdong, China

Ying Jin School of Automation, Nanjing University of Science and Technology, Nanjing, China

Ma Jingjing College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China

Liwei Ju School of Economics and Management, North China Electric Power University, Beijing, China

Ling Kang Department of Aeronautics, School of Physics and Mechanical & Electrical Engineering, Xiamen University, Xiamen, China

Qiyan Ke The School of Information Engineering, Wuhan University of Technology, Wuhan, China

Jinjiao Kong China Electric Power Research Institute, Beijing, China

Lei Lan College of Electrical Engineering, Wuhan University, Wuhan, China

Xufeng Lang State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing, China

Cheng Le Department of Computer Science and Engineering, Huaian College of Information Technology, Huaian, China

Jian Le School of Electrical Engineering, Wuhan University, Wuhan, China

Min Lei Institute of Information and Control Technology, North China Institute of Science and Technology, Beijing, China

Baogang Li Naval Aeronautical and Astronautical University, Yantai, China

Chongjian Li Automation Research and Design Institute of Metallurgical Industry, Beijing, China

Dajian Li No. 365 Research Institute, Northwestern Polytechnical University, Xi'an, China

Dan Li North China Electric Power University, BeiJing, China

Dongmei Li Key Laboratory of Dynamics and Control of Flight Vehicle, Ministry of Education, School of Aerospace Engineering, Beijing Institute of Technology, Beijing, China

Duanchao Li Dispatching and Control Center, Anhui Electric Power Corporation, Hefei, China

Fang Li The State Key Laboratory of Tribology, Tsinghua University, Beijing, China

Guolin Li Naval Aeronautics and Astronautics University, Yantai, China

Haijun Li Naval Aeronautical and Astronautical University, Yantai, China

Hang Li School of Control and Computer Engineering, North China Electric Power University, Beijing, China

Jianwen Li School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Jiayu Li School of Economics and Management, North China Electric Power University, Beijing, China

Jingchun Li State Radio Monitoring Center, Beijing, China

Junlou Li Nanjing Institute of Tourism and Hospitality, Nanjing, China

Junqing Li School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Kejie Li School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China

Ming Li College of Electrical Engineering & New Energy, Three Gorges University, Yichang, China

Mingjun Li Department of Urban Rail Transportation, Beijing Vocational College of Transportation, Beijing, China

Nengshao Li Department of Electronic and Information Engineering, Inner Mongolia University, Hohhot, China

Pingkang Li School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, China

Quanhu Li Department of Electronic and Information Engineering, Inner Mongolia University, Hohhot, China

Rui Li School of Automation and Electrical and Engineering, University of Science and Technology Beijing(USTB), Beijing, China

College of Electrical Engineering, Wuhan University, Wuhan, China

Ruochen Li School of Economics and Management, North China Electric Power University, Beijing, China

Shu Li BeiHang University, Beijing, China

Tong Li College of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Wei Li College of Electronic Engineering, Wuhan University, Wuhan, China

Wenjing Li Chongqing Water Resources and Electric Engineering College, Chongqing, China

Xinli Li School of Control and Computer Engineering, North China Electric Power University, Beijing, China

Yanan Li State Power Economic Research Institute, Beijing, China

Yanlong Li Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, Wuhan, China

Yeli Li The Computer Science and Technology Department, Beijing Institute of Graphic Communication, Beijing, China

Yingying Li College of Optoelectronic Science and Engineering, National University of Defense Technology, Changsha, China

Yitao Li Hangzhou Dahe Thermo-magnetics Co. Ltd., Hangzhou, China

Yiwei Li School of Electrical and Electronic, Engineering North China Electric Power University, Beijing, China

Yonggang Li School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Yuan Li School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China

Zhongxing Li Guangzhou Academy of Special Equipment Inspection & Testing, Guangzhou, China

Jianwei Lian Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, Wuhan, China

Jing Liang School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan, China

Qingyun Liang College of Electrical Engineering & New Energy, Three Gorges University, Yichang, China

Xiao Liang School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, China

Chuanglu Lin Guangzhou Academy of Special Equipment Inspection & Testing, Guangzhou, China

Gaoxiang Lin Wenzhou Power Supply Company, Wenzhou, China

Hui Lin Department of Aeronautics, School of Physics and Mechanical & Electrical Engineering, Xiamen University, Xiamen, China

Jianzhi Lin Qingyuan Power Supply Company, Qingyuan, China

Huanzhang Ling College of Automation, Harbin Engineering University, Harbin, China

Changsheng Liu School of Automation, Nanjing University of Science and Technology, Nanjing, China

Chao Liu Institute of Systems Science, Northeastern University, Shenyang, China

Haibo Liu School of Economics and Management, North China Electric Power University, Beijing, China

Huijia Liu Electrical Engineering & Renewable Energy School, China Three Gorges University, Yichang, China

Huixia Liu No. 365 Research Institute, Northwestern Polytechnical University, Xi'an, China

School of Automation, Northwestern Polytechnical University, Xi'an, China

Kaipei Liu School of Electrical Engineering, Wuhan University, Wuhan, China

Limin Liu College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China

Ming Liu Science College, Air Force Engineering University, Xi'an, China

Minxia Liu School of Electronic Engineering, Dongguan University of Technology, Guangdong, China

Peiyong Liu College of Life and Health Sciences, Northeastern University, Shenyang, China

Qianjin Liu School of Electric Power, South China University of Technology, Guangzhou, China

Qifeng Liu Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Centre, Wuhan, China

Si Liu North China Electric Power University, Beijing, China

Xiang Liu School of management, Shanghai University, Shanghai, China

Xijia Liu Wenzhou Power Supply Company, Wenzhou, China

Yang Liu Shenyang Institute of Automation Chinese Academy of Sciences, Shenyang, China

Shenyang University, Shenyang, China

University of Chinese Academy of Sciences, Beijing, China

Yansheng Liu 95997 Troops of Chinese PLA, Beijing, China

Yijing Liu Science College, Air Force Engineering University, Xi'an, China

Yongchun Liu Beijing Institue of Technology, Beijing, China

Yu Liu College of Electronic Engineering, Wuhan University, Wuhan, China

Yuan Liu College of Electrical Engineering & New Energy, China Three Gorges University, Yichang, China

Yueting Liu School of Electronics and Information Engineering, Lanzhou University of Arts and Science, Lanzhou, China

Zhihao Liu Naval Submarine Academy, Qingdao, China

Fei Long The 712 Research Institute, China Shipbuilding Industry Corporation, Wuhan, China

Cuihua Lu Naval Aeronautics and Astronautics University, Yantai, China

Hailiang Lu College of Electrical Engineering, Wuhan University, Wuhan, China

Wei Lu The School of Information Engineering, Wuhan University of Technology, Wuhan, China

Feng Luo Clean Energy Automotive Engineering Center, School of Automotive Studies, Tongji University, Shanghai, China

Hongbo Luo College of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou Higher Education Mega Center, Guangzhou, China

Tinade Luo Tobacco Guangdong Industrial Co., Ltd., Guangzhou, China

Yuesheng Luo College of Automation, Harbin Engineering University, Harbin, China

Jie Lv College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China

Xiaofeng Lv Naval Aeronautical and Astronautical University, Yantai, China

Jian Ma Laser Research Institute, Shandong Academy of Science, Jinan, China

Jinhui Ma Dispatching and Control Center, Anhui Electric Power Corporation, Hefei, China

Ling Ma Naval Aeronautical and Astronautical University, Yantai, China

Wei Ma State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing, China

Nanjing Institute of Tourism and Hospitality, Nanjing, China

Yiping Ma College of Mechanical and Electrical, Changchun University of Science and Technology University, Changchun, China

Zhi q iang Ma Communication Training Base of the PLA General Staff Headquarters, XuanHua, China

Zhiqiang Ma College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China

Kai Mao State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Zhejiang, China

Xiaoqing Mei Shanghai IoT Co., LTD., Shanghai Research Center for Wireless Communications, Shanghai, China

Gangyang Nan Laser Research Institute, Shandong Academy of Science, Jinan, Chinas

Yimin Nie College of Resources and Environment, Shandong Agricultural University, Tai'an, China

Guobiao Ning Qingyuan Power Supply Company, Qingyuan, China

Shuang Pan Naval Submarine Academy, Qingdao, China

Wei Pang Naval Aeronautical and Astronautical University, Yantai, China

Shirui Peng Air Force Early Warning Academy, Wuhan, China

Tao Peng School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan, China

Wenhui Peng Navy Submarine Academy, Qingdao, China

Tran Ky Phuc Vietnam Institute of Energy, Hanoi, Vietnam

Bibo Ping Science and Technology on Electromagnetic Compatibility Laboratory, Wuhan, China

Xiandeng Ping School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, China

Yan Pu Sichuan Institute of Solid-State Circuits, Chongqing, China

Yali Qi School of Electrical Engineering and Automation, Tianjin University, Tianjin, China

The Computer Science and Technology Department, Beijing Institute of Graphic Communication, Beijing, China

Bifu Qian Wenzhou Power Supply Company, Wenzhou, China

Guohua Qiao State Grid Hebei Electric Power Company, Shijiazhuang, China

Jing Qiao Tobacco Guangdong Industrial Co., Ltd., Guangzhou, China

Chao Qin School of Economics and Management, North China Electric Power University, Beijing, China

Mengtao Qin Science and Technology on Electromagnetic Compatibility Laboratory, Wuhan, China

Xiao Qin Department of Computer Science and Software Engineering, Auburn University, Auburn, AL, USA

Hongkai Qiu School of Control and Computer Engineering, North China Electric Power University, Beijing, China

Runhe Qiu College of Information Sciences and Technology, Engineering Research Center of Digitized Textile & Fashion Technology, Ministry of Education, Donghua University, Shanghai, China

Lu Qu College of Electrical Engineering, Wuhan University, Wuhan, China

Xiaoyan Qu Naval Aeronautical and Astronautical University, Yantai, China

Jiangtao Quan Electric Power Research Institute of Hubei Power Grid Corporation, Wuhan, China

Dashun Que Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education, Wuhan University of Technology, Wuhan, China

Guyin Rao College of Optoelectronic Science and Engineering, National University of Defense Technology, Changsha, China

Ping Ren Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, Wuhan, China

Tianpeng Ren National Key Laboratory of Science and Technology on Aerospace Flight Dynamics, Beijing Aerospace Control Center, Beijing, China

Zheng Ren College of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Lianhai Shan Shanghai IoT Co., LTD., Shanghai Research Center for Wireless Communications, Shanghai, China

Yuhao Shan Navy Submarine Academy, Qingdao, China

Haiyan Shao School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China

School of Mechanical engineering, University of Jinan, Jinan, China

Haiyan Shao School of Mechatronical Engineering, Beijing Institute of Technology, Beijing, China

Jiejie Shao State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Zhejiang, China

Fengwen Shen College of Electrical and Information Engineering, Hunan University, Changsha, China

Tianfu Shen Electric Power of South China University of Technology, GuangZhou, China

Tianfu Shen Electric Power of South China University of Technology, Guangzhou, Guangdong, China

Xu Shen Thermal Power Plant, Daqing Petro-chemical Branch Company, Daqing, China

Yong Shen 95997 Troops of Chinese PLA, Beijing, China

Chao Shi School of Electric Power, South China University of Technology, Guangzhou, China

Jianli Shi Navy Submarine Academy, Qingdao, China

Wenbo Shi Electric Power of South China University of Technology, Guangzhou, Guangdong, China

Xiaoqiao Shi The School of Electro-Mechanical Engineering, Xidian University, Xi'an, China

Yujie Shi School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China

Jiangfeng Song Laser Research Institute, Shandong Academy of Science, Jinan, China

Jingming Song Institute of Optoelectronic Technology, Beijing University of Aeronautics and Astronautics, Beijing, China

Lili Song School of Information Engineering, Inner Mongolia University of Technology, Huhhot, China

Mofei Song State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing, China

Ningfang Song Institute of Optoelectronic Technology, Beijing University of Aeronautics and Astronautics, Beijing, China

Wenwu Song Science and Technology on Electromagnetic Compatibility Laboratory, Wuhan, China

Chengzhi Su College of Mechanical and Electrical, Changchun University of Science and Technology University, Changchun, China

Xujun Su Mechanical Engineering College, Shijiazhuang, China

Jihua Sun Laser Research Institute, Shandong Academy of Science, Jinan, China

Jing Sun Naval Aeronautical and Astronautical University, Yantai, China

Xiaofang Sun Electrical Engineering & Renewable Energy School, China Three Gorges University, Yichang, China

Xiaohong Sun National ASIC System Engineering Research Center, Southeast University, Nanjing, China

Xiaolei Sun Institute of Information and Control Technology, North China Institute of Science and Technology, Beijing, China

Yukun Sun Nanjing Institute of technology, Nanjing, China

Zhengxing Sun State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing, China

Jin Tan Guangdong Electric Power Research Institute, Guangzhou, China

Zhongfu Tan School of Economics and Management, North China Electric Power University, Beijing, China

Geishi Tang National Key Laboratory of Science and Technology on Aerospace Flight Dynamics, Beijing Aerospace Control Center, Beijing, China

Jing Teng School of Control and Computer Engineering, North China Electric Power University, Beijing, China

Zhaosheng Teng College of Electrical and Information Engineering, Hunan University, Changsha, China

Vu Toan Thang Electricity Regulatory Authority of Vietnam, Hanoi, Vietnam

Chaonan Tong School of Automation and Electrical and Engineering, University of Science and Technology Beijing(USTB), Beijing, China

Jie Tong China Electric Power Research Institute, Beijing, China

Xuefang Tong China Electric Power Research Institute, Wuhan, China

Jianxiong Wan College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China

Lili Wan School of Automation, Huazhong University of Science and Technology, Wuhan, China

Chengjiang Wang College of Electrical Engineering & New Energy, Three Gorges University, Yichang, China

Chengsheng Wang Automation Research and Design Institute of Metallurgical Industry, Beijing, China

Demin Wang College of Mechanical and Electrical, Changchun University of Science and Technology University, Changchun, China

Dong Wang School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Feng Wang China Electric Power Research Institute, Beijing, China

Suzhou Innotion Tech. Co., Ltd., Suzhou, China

Guan Wang School of Economics and Management, North China Electric Power University, Beijing, China

Guangxue Wang Air Force Early Warning Academy, Wuhan, China

Hong Wang Chengdu Neusoft university, Chengdu, China

Hongqiang Wang 95997 Troops of Chinese PLA, Beijing, China

Hua Wang School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing, China

Jian Wang Electric Power of South China University of Technology, GuangZhou, China

Jian Wang School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China

Jian Wang School of Mechatronical Engineering, Beijing Institute of Technology, Beijing, China

Jianping Wang School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, China

Jiaxing Wang Wenzhou Power Supply Company, Wenzhou, China

Jinhao Wang State Grid Shanxi Electric Power Research Institute, Taiyuan, China

Juan Wang Wuhan No.2 Light Industry School, Wuhan, China

Li Wang School of Electrical Engineering, Wuhan University, Wuhan, China

Liguo Wang JiangMen Power Bureau, Jiangmen, China

Lingyun Wang College of Electrical Engineering & New Energy, China Three Gorges University, Yichang, China

Min Wang School of Automation, Huazhong University of Science and Technology, Wuhan, China

Qiwu Wang Laser Research Institute, Shandong Academy of Science, Jinan, China

Rongtao Wang Wenzhou Power Supply Company, Wenzhou, China

Tian Wang College of Information Sciences and Technology, Engineering Research Center of Digitized Textile & Fashion Technology, Ministry of Education, Donghua University, Shanghai, China

Wei Wang Mechanical Engineering College, Shijiazhuang, China

Ximei Wang School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Xindi Wang School of Electronic Power Engineering, Nanjing Institute of Technology, Nanjing, China

Xing q uan Wang Communication Training Base of the PLA General Staff Headquarters, XuanHua, China

Xinghai Wang Naval Aeronautical and Astronautical University, Yantai, China

Xinhua Wang Guangzhou Academy of Special Equipment Inspection & Testing, Guangzhou, China

Y ujian Wang State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing, China

Yanjing Wang College of Mechanical and Electrical, Changchun University of Science and Technology University, Changchun, China

Yixuan Wang Lanzhou Jiaotong University, Lanzhou, China

Yu Wang College of Electrical Engineering, Wuhan University, Wuhan, China

Zhenyu Wang School of Control and Computer Engineering, North China Electric Power University, Beijing, China

Zhongyan Wang Sichuan Institute of Solid-State Circuits, Chongqing, China

Jian Wei College of Electrical Engineering, Wuhan University, Wuhan, China

Congxi Wen North China Electric Power University, Beijing, China

He Wen College of Electrical and Information Engineering, Hunan University, Changsha, China

Xishan Wen College of Electrical Engineering, Wuhan University, Wuhan, China

Xishan Wen School of Electrical Engineering, Wuhan University, Wuhan, China

Dong Wu School of Automation, Beijing Institute of Technology, Beijing, China

Fengtao Wu Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Centre, Wuhan, China

Jianhui Wu National ASIC System Engineering Research Center, Southeast University, Nanjing, China

Jie Wu School of Electrical Engineering, Wuhan University, Wuhan, China

Li z hen Wu College of Computer and Communcation, Lanzhou University of Technology, Lan zhou, China

Nan Wu Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, Wuhan, China

Qi Wu Electronic Science and Engineering Department, National University of Defense Technology, Changsha, China

Shihong Wu College of Electrical Engineering, Wuhan University, Wuhan, China

Shipu Wu Electric Power Research Institute of China, Wuhan, China

Xingjun Wu Guangzhou Academy of Special Equipment Inspection & Testing, Guangzhou, China

Xixiu Wu College of Automation, Wuhan University of Technology, Wuhan, China

Qingbiao Xi No. 365 Research Institute, Northwestern Polytechnical University, Xi'an, China

School of Automation, Northwestern Polytechnical University, Xi'an, China

Tan Xia Wuhan Vocational College of Software and Engineering, Wuhan, China

Qing Xie College of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Yunyun Xie School of Automation, Nanjing University of Science and Technology, Nanjing, China

Yong Xiong Shanghai Institute of Microsystem and Information Technology, CAS, Shanghai, China

Chengdong Xu Key Laboratory of Dynamics and Control of Flight Vehicle, Ministry of Education, School of Aerospace Engineering, Beijing Institute of Technology, Beijing, China

Huiming Xu School of Electric Power, South China University of Technology, Guangzhou, China

Man Xu College of Science, Harbin Engineering University, Harbin, China

Mingyan Xu Institute of Information and Control Technology, North China Institute of Science and Technology, Beijing, China

Ouyang Xu State Grid GanSu Maintenance Company, Lanzhou, China

Tao Xu School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China

Ting Xu Department of Electronic and Information Engineering, Inner Mongolia University, Hohhot, China

Yang Xu College of Electrical Engineering, Wuhan University, Wuhan, China

Yonghai Xu North China Electric Power University, Beijing, China

Yuqin Xu College of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Chunyu Yan China Electric Power Research Institute, Beijing, China

Haiyan Yan Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, Wuhan, China

Can Yang College of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China

Chaoying Yang State Grid Shanxi Electric Power Company, Taiyuan, China

Guojin Yang Economic Research Institute of Qinghai Province, Xining, China

Hao Yang School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Hu Yang Electronic Science and Engineering Department, National University of Defense Technology, Changsha, China

Hui Yang 95997 Troops of Chinese PLA, Beijing, China

Jie Yang School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing, China

Juan Yang School of Information and Technology, Jinling Institute of Technology, Nanjing, China

Li Yang Department of Mathematics, Liaoning University, Shenyang, China

Yong Yang Gansu Electric Power Research Institute, Lanzhou, China

Yujie Yang Institute of Optoelectronic Technology, Beijing University of Aeronautics and Astronautics, Beijing, China

Zhifei Yang Lanzhou Jiaotong University, Lanzhou, China

Chengzhu Yao Naval Aeronautical and Astronautical University, Yantai, China

Wei Yao Shanghai IoT Co., LTD., Shanghai Research Center for Wireless Communications, Shanghai, China

Peiqing Ye The State Key Laboratory of Tribology, Tsinghua University, Beijing, China

Peng Ye Tancy Group Co. Ltd, Zhejiang, China

Wen Ye Naval Aeronautical and Astronautical University, Yantai, China

Nie Yongfang Naval Submarine Academy, Qingdao, China

Lin You College of Communication Engineering, Hangzhou Dianzi University, Hangzhou, China

Feng Yu School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan, China

Haibo Yu School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Tao Yu Electric Power College, South China University of Technology, Guangzhou, China

Jie Yuan College of Optoelectronic Science and Engineering, National University of Defense Technology, Changsha, China

Weiyi Zeng Huazhong University of Science and Technology, Wuhan, China

Dehui Zhai School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China

Mingyue Zhai North China Electric Power University, BeiJing, China

Yanwei Zhai Clean Energy Automotive Engineering Center, School of Automotive Studies, Tongji University, Shanghai, China

Wei Zhan Army 71217 of PLA, Shandong, China

Xiang Zhan College of Optoelectronic Science and Engineering, National University of Defense Technology, Changsha, China

Xiangling Zhan College of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Dongfu Zhang Shanghai Regeon Electric Co., Ltd., Shanghai, China

Guoshan Zhang School of Electrical Engineering and Automation, Tianjin University, Tianjin, China

Haibo Zhang College of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China

Jian Zhang Department of Urban Rail Transportation, Beijing Vocational College of Transportation, Beijing, China

Jingzhuo Zhang College of Electronic Engineering, Wuhan University, Wuhan, China

Kai Zhang China Ship Development and Design Center, Wuhan, China

Lanlan Zhang Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China

Lei Zhang College of Computer and Communcation, Lanzhou University of Technology, Lan zhou, China

Liang Zhang Nanjing Institute of technology, Nanjing, China

School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China

Linghao Zhang School of Automation, Nanjing University of Science and Technology, Nanjing, China

Lipeng Zhang School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China

Niaona Zhang Changchun University of Technology, Changchun, China

Quanzhu Zhang Institute of Information and Control Technology, North China Institute of Science and Technology, Beijing, China

Tianyu Zhang Department of Aeronautics, School of Physics and Mechanical & Electrical Engineering, Xiamen University, Xiamen, China

Wei Zhang Science and Technology on Electromagnetic Compatibility Laboratory, Wuhan, China

XiuXia Zhang School of Electronics and Information Engineering, North National University, Yinchuan, China

Yong Zhang The State Key Laboratory of Tribology, Tsinghua University, Beijing, China

Yongliang Zhang Naval Aeronautics and Astronautics University, Yantai, China

Zhe Zhang Changchun University of Technology, Changchun, China

Zhenhai Zhang School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China

Zhenhai Zhang School of Mechatronical Engineering, Beijing Institute of Technology, Beijing, China

Zhongyuan Zhang Gansu Electric Power Corporation, Lanzhou, China

Junlong Zhao North China University of Water Resources and Electric Power, Zhengzhou, China

Minrong Zhao Science College, Air Force Engineering University, Xi'an, China

Xiongwen Zhao School of Electrical and Electronic, Engineering North China Electric Power University, Beijing, China

Yanhong Zhao North China Electric Power University, BeiJing, China

Jian g uo Zheng Communication Training Base of the PLA General Staff Headquarters, XuanHua, China

Jingjing Zheng Gansu Electric Power Research Institute, Lanzhou, China

Shiling Zheng Anhui Electric Power Research Institute, Hefei, China

Bin Zhou Electric Power College, South China University of Technology, Guangzhou, China

Sujuan Zhou College of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou Higher Education Mega Center, Guangzhou, China

Xuanqing Zhou College of Electrical Engineering & New Energy, China Three Gorges University, Yichang, China

Yabing Zhou Qingyuan Power Supply Company, Qingyuan, China

Yunfei Zhou School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China

Changcheng Zhu Electric Power Research Institute of Hubei Power Grid Corporation, Wuhan, China

Chunyi Zhu Automation Research and Design Institute of Metallurgical Industry, Beijing, China

Jian Zhu Key Laboratory of Network Coding Key Technology and Application, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China

Qin Zhu School of Electrical and Information Technology, Yunnan University of Nationalities, Kunming, China

Yaochun Zhu Longyuan (Beijing) Wind Power Engineering Technology Co., Ltd, Beijing, China

Yunlong Zhu Shenyang Institute of Automation Chinese Academy of Sciences, Shenyang, China

Ming Zong School of Electrical Engineering, Shenyang University of Technology, Shenyang, China

Sanyong Zou Tinghua University, Beijing, China

Hao Zuo Beijing Institute of Technology, Beijing, China

Jiajun Zuo Air Force Early Warning Academy, Wuhan, China

Part I
Algorithm Design

Chapter 1
Modeling and Simulation of the Emergency Diesel Generator Based on MATLAB

Dongfu Zhang and Jianhua Cui

Abstract The loading calculation is one of the key technologies for the emergency diesel generator which plays an important role in nuclear power. For the purpose of analyzing the process of loading, this chapter has established a simulation model of the emergency diesel generator based on MATLAB/Simulink, which may accurately reflect the mechanical and electromagnetic coupling process and the nature of the strong coupling and nonlinearity. The author simulates the loading process of the emergency diesel generator with the simulation results showing that the model is correct and reasonable; in the meanwhile, a theoretical guidance for the capacity choices of the emergency diesel generator is provided.

Keywords Diesel generators · Simulation · Loading calculation

1.1 Introduction

In the nuclear power plants, it is important to ensure the continuity of power supply. The emergency power system must be adopted except for the redundant power supply line. The world widely uses the emergency diesel generator (EDG) to provide emergency power for the emergency bus in the loss of normal power supply and thus ensures continuous power supply for the critical equipment.

Under the background of nuclear localization, it is significant to analyze and simulate the EDG, which has not been localized yet. The simulation analysis of asynchronous motor loading process is a key technology in EDG that determines the selection of diesel and generator and the overall design requirements of EDG.

The EDG is a strongly coupled and complex nonlinear system. The author establishes a detailed mathematical model of the EDG and analyses the process of loading induction motors.

D. Zhang (✉) · J. Cui
Shanghai Regeon Electric Co., Ltd., 201612 Shanghai, China
e-mail: zdf403217@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_1

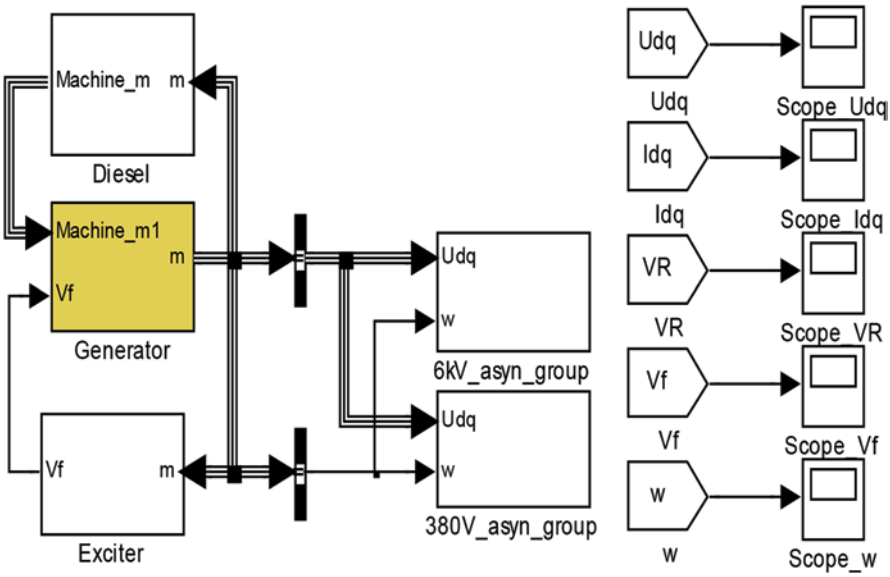


Fig. 1.1 Model of an emergency diesel generator

1.2 Mathematical Model of the Emergency Diesel Generator

The EDG mathematical model, which is established based on the MATLAB/Simulink platform under the guidance of modular thinking, is shown in Fig. 1.1. The synchronous generator model, excitation system model, diesel, and its speed controller model and asynchronous motor model are included in the model.

1.2.1 Mathematical Model of the Synchronous Generator

The generator with uniform air gap can be considered as a special case of nonuniform air-gap generator; therefore, it is universal to analyze the salient rotor generator. The author focuses on the analysis of the salient rotor generator's mathematical model.

The electromechanical transient model is sufficient to characterize the generator features [1]; thus, the five-order practical generator model is adopted:

$$\begin{cases} u_{\mathrm{d}} = E_{\mathrm{d}}'' + X_{\mathrm{q}}'' i_{\mathrm{q}} - R_{\mathrm{s}} i_{\mathrm{d}} \\ u_{\mathrm{q}} = E_{\mathrm{q}}'' - X_{\mathrm{d}}'' i_{\mathrm{d}} - R_{\mathrm{s}} i_{\mathrm{q}} \end{cases}.$$

$$\begin{cases} T_{\mathrm{d0}}' pE_{\mathrm{q}}' = E_{\mathrm{f}} - (E_{\mathrm{q}}' - X_{\mathrm{d}\tau} E_{\mathrm{q}}' + X_{\mathrm{d}\tau} E_{\mathrm{q}}'') \\ T_{\mathrm{d0}}'' pE_{\mathrm{q}}'' = -E_{\mathrm{q}}'' + E_{\mathrm{q}}' - (X_{\mathrm{d}}' - X_{\mathrm{d}}'') i_{\mathrm{d}} \\ T_{\mathrm{q0}}'' pE_{\mathrm{d}}'' = -E_{\mathrm{d}}'' + (X_{\mathrm{q}} - X_{\mathrm{q}}'') i_{\mathrm{q}} \end{cases}. \tag{1.1}$$

$$\begin{cases} T_{\mathrm{J}} p\omega = T_{\mathrm{m}} - [E_{\mathrm{q}}'' i_{\mathrm{q}} + E_{\mathrm{d}}'' i_{\mathrm{d}} + \left(X_{\mathrm{d}}'' - X_{\mathrm{q}}''\right) i_{\mathrm{d}} i_{\mathrm{q}}] - D\omega \\ p\delta = \omega - 1 \end{cases},$$

wherein p is the differential operator, E_{d}'' is the d-axis sub-transient force, E_{q}'' is the q-axis sub-transient force, E_{q}' is the q-axis transient force, E_{f} is the stator excitation force; T_{m} is the drag torque of the generator, and D is the generator damping [2].

1.2.2 Mathematical Model of the Excitation System

Most of the EDG excitation systems use the alternating current (AC) exciter. AC exciter with rotary rectifier, from which carbon brushes and slip rings are removed, is more reliable and highlights higher performance than the exciter with a static rectifier.

The author has established a mathematical model of an AC exciter that takes into account the effect of magnetic saturation, voltage drop caused by changing arc, and demagnetization caused by armature reaction. The schematic diagram is shown in Fig. 1.2.

1.2.2.1 Saturation Coefficient of an AC Exciter

Figures 1.3 and 1.4 show the excitation characteristics of the AC exciter:

$$\frac{E_{\mathrm{FD}}}{V_{\mathrm{R}}} = \frac{K/(R(1+R/r))}{\left(\dfrac{1-K/r}{1+R/r}\right) + \dfrac{M}{R}s} \tag{1.2}$$

$$V_{\mathrm{E}} = E_{\mathrm{FD}} = KI_{\mathrm{FF}}(B/A), \tag{1.3}$$

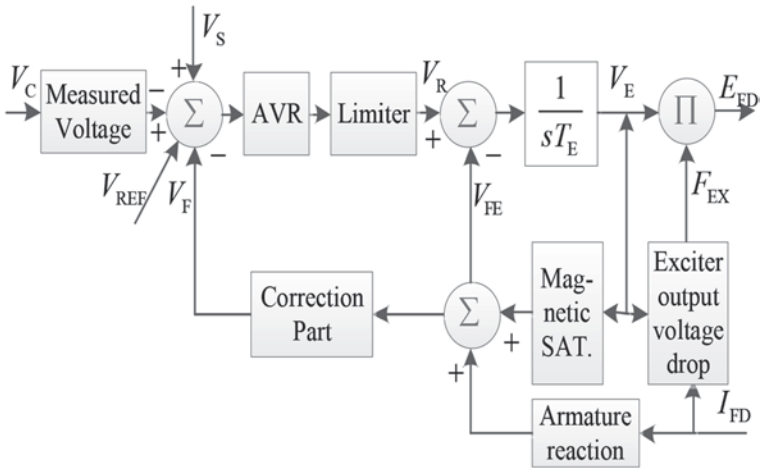


Fig. 1.2 Schematic diagram of an AC exciter. *AVR* automatic voltage regulator

wherein E_{FD} is the exciter output voltage, V_R is the excitation voltage of the exciter, and the meaning of the other parameters above is shown in Figs. 1.3 and 1.4 [3].

Combining Eqs. (1.2) and (1.3), we can get the following relationship, which takes the magnetic saturation into account:

$$\frac{E_{FD}}{V_R} = \frac{K/(R(1+R/r))}{K_E + ((A-B)/B)/(1+R/r) + T_E s} = \frac{K/(R(1+R/r))}{K_E + S_E + T_E s}, \tag{1.4}$$

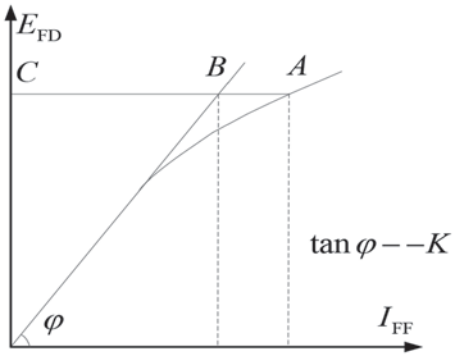


Fig. 1.3 Exciter saturation characteristics

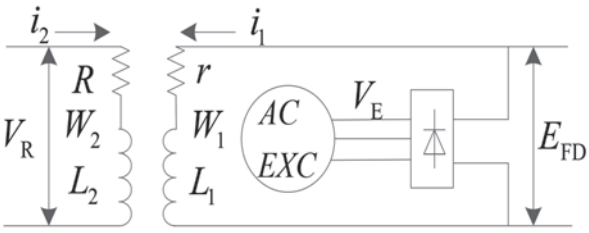


Fig. 1.4 Exciter structure. *AC* alternating current

wherein $S_E = ((A-B)/B)/(1+R/r)$, which is the exciter saturation factor, and K_E is the self-excited coefficient of the AC exciter. If we excite the exciter separated, r equals ∞, and S_E equals $(A-B)/B$.

1.2.2.2 Armature Reaction of the AC Exciter

The biggest difference between the AC and the direct current (DC) exciter on the characteristics is that the armature reaction of the AC exciter is strong and must be taken into consideration. The load current of the AC exciter can be considered as a vector on the vertical axis; and it uses K_D to represent the armature reaction coefficient.

1.2.2.3 Voltage Drop Caused by Changing Arc

When the arc is changed, the voltage drop will be produced. If the voltage drop of rectifiers is ignored, the DC output voltage equals to no-load voltage minus the voltage drop caused by changing arc. According to the current operating characteristics of the three-phase rectifier bridge, we can derive the corresponding formula [4]:

$$\begin{cases} E_{FD} = V_E(1-0.577I_N), \text{when } I_N \le 0.433 \\ E_{FD} = V_E\left(\sqrt{0.75-{I_N}^2}\right), \text{when } 0.433 < I_N < 0.75 \\ E_{FD} = V_E\left(\sqrt{3}(1-I_N)\right), \text{when } 0.75 \le I_N \le 1 \end{cases}. \tag{1.5}$$

1.2.2.4 Mathematical Model of the Automatic Voltage Regulator (AVR)

The voltage regulator mainly consists of the voltage detecting part, the comparing part, and the correction part. The voltage detection part can be generally expressed by the first-order inertial link with its transfer function as:

$$W_B(s) = \frac{K_T}{1+T_T s}, \tag{1.6}$$

where K_T is the ratio coefficient of the RC filter and T_T is the time constant.

In order to improve the static and dynamic characteristics of the excitation system, it is necessary to correct the excitation system. The transfer function of correction link is:

$$W_F(s) = \frac{K_F s}{1+T_F s}, \tag{1.7}$$

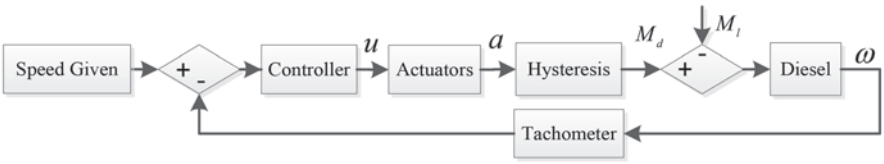


Fig. 1.5 The structure of diesel and its governor

where K_F and T_F are the ratio coefficient and time constant of the feedback, respectively.

1.2.3 Mathematical Model of the Diesel Engine and Its Governor

The diesel engine and its governor are a complex, nonlinear, and dynamic system, whose structure diagram is shown in Fig. 1.5.

1.2.3.1 Mathematical Model of the Controller

The controller is mainly used to adjust the speed of the diesel engine, while the proportional integral derivative (PID) controller is widely adopted in engineering. The author tunes the parameters of the controller with the critical ratio method and obtains satisfactory control performance. The speed measurement is generally expressed by the first-order inertia link. If its dynamic impact on the governor is ignored, its transfer function can be set to 1.

1.2.3.2 Mathematical Model of the Actuator

The actuator's role is to convert the controller output to the axial displacement, which is positively correlated with the controller output. The actuator transfer function is:

$$\frac{a(s)}{u_c(s)} = \frac{K_1}{T_1 s + 1}, \tag{1.8}$$

wherein K_1 is the actuator gain and T_1 is the time constant of the actuator.

1.2.3.3 Mathematical Model of Torque Lags

As to certain diesel engine, the time delay caused by actuator, combustion process, and thermodynamic process can be expressed by the following formula:

$$\frac{M_{\mathrm{d}}(s)}{a(s)}=\frac{1}{t_{\mathrm{d}}s+1}, \tag{1.9}$$

where t_{d} is the time constant, which can be calculated according to the formula: $(15/n_{\mathrm{n}})\le t_{\mathrm{d}}\le(15/n_{\mathrm{n}}+120/(n_{\mathrm{n}}i))$, where n_{n} represents the rated speed of diesel engine and i represents the number of strokes of diesel engine.

1.2.3.4 Mathematical Models of Diesel

The diesel engine normally can be expressed by first-order transfer function:

$$\frac{\omega(s)}{M_{\mathrm{d}}(s)-M_{\mathrm{l}}(s)}=\frac{1}{T_{\mathrm{a}}s+C_{\mathrm{d}}}, \tag{1.10}$$

wherein C_{d} and T_{a} refer to the damping coefficient and the time constant of the diesel engine, respectively.

1.2.4 Mathematical Model of the Induction Motor

In general, the stator winding transient process of asynchronous motors is much faster than the electromagnetic transient process of rotor winding, and even faster than the power system transient process [5]; therefore, the stator transient process of induction motor can be ignored. The Park equation of induction motor is listed as below:

$$\begin{bmatrix}u_{\mathrm{ds}}\\u_{\mathrm{qs}}\\0\\0\end{bmatrix}=p\begin{bmatrix}0\\0\\\Psi_{\mathrm{dr}}\\\Psi_{\mathrm{qr}}\end{bmatrix}+\begin{bmatrix}-\omega\Psi_{\mathrm{qs}}\\\omega\Psi_{\mathrm{ds}}\\-(\omega-\omega_{\mathrm{r}})\Psi_{\mathrm{qr}}\\(\omega-\omega_{\mathrm{r}})\Psi_{\mathrm{dr}}\end{bmatrix}+\begin{bmatrix}r_{\mathrm{s}}&0&0&0\\0&r_{\mathrm{s}}&0&0\\0&0&r_{\mathrm{r}}&0\\0&0&0&r_{\mathrm{r}}\end{bmatrix}\begin{bmatrix}i_{\mathrm{ds}}\\i_{\mathrm{qs}}\\i_{\mathrm{dr}}\\i_{\mathrm{qr}}\end{bmatrix}$$

$$\begin{bmatrix}\Psi_{\mathrm{ds}}\\\Psi_{\mathrm{qs}}\\\Psi_{\mathrm{dr}}\\\Psi_{\mathrm{qr}}\end{bmatrix}=\begin{bmatrix}L_{\mathrm{s}}&0&L_{\mathrm{m}}&0\\0&L_{\mathrm{s}}&0&L_{\mathrm{m}}\\L_{\mathrm{m}}&0&L_{\mathrm{r}}&0\\0&L_{\mathrm{m}}&0&L_{\mathrm{r}}\end{bmatrix}\begin{bmatrix}i_{\mathrm{ds}}\\i_{\mathrm{qs}}\\i_{\mathrm{dr}}\\i_{\mathrm{qr}}\end{bmatrix} \tag{1.11}$$

$$\begin{cases}T_{\mathrm{e}}=\dfrac{3p_{\mathrm{j}}}{2}L_{\mathrm{m}}(i_{\mathrm{qs}}i_{\mathrm{dr}}-i_{\mathrm{ds}}i_{\mathrm{qr}})\\\dfrac{d\omega_{\mathrm{r}}}{dt}=\dfrac{p_{\mathrm{j}}}{2J}(T_{\mathrm{e}}-T_{\mathrm{l}})\end{cases},$$

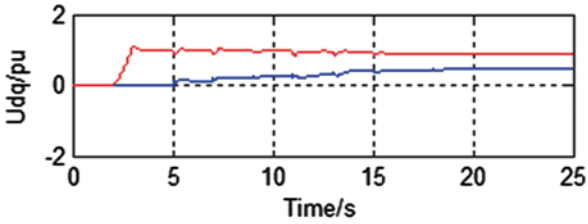


Fig. 1.6 Terminal voltage curve

where L_{ls}, L_{lr}, L_s, L_r, L_m,p_j, and J refer to the stator leakage inductance, the rotor leakage inductance, the stator inductance, rotor inductance, the mutual inductance between stator and rotor, the number of pole pairs, and the moment of inertia, respectively. T_e, T_l, ω_r refer to the motor electromagnetic torque, the motor load torque, and the instantaneous angular velocity of the motor rotor, respectively.

1.3 Simulation Results

Choose one synchronous generator which is rated at 5MVA with the rated line voltage of 6 kV and rated frequency of 50 Hz.

Main parameters of the diesel engine: t_d is 0.042, T_a is 2.247, C_d takes 0.03. PID controller parameters: K_P is 20, T_I is 0.1, T_D takes 0.05.

Model *6kV_asyn_group* and model *380V_asyn_group*, in the simulation model, used to simulate the EDG loading process, shown in Fig. 1.1, and are used to simulate the plant pump, the safety injection pump, the spray pump, the feed pump, the chiller, the blower, etc.

The EDG automatically starts, accelerates, and establishes voltage when the security segment loses power supply. If the entire load is loaded at the same time, the voltage and frequency will fluctuate drastically, and the requirement for the capacity of EDG is relatively high; therefore, the EDG loads the asynchronous motor step by step that would reduce the capacity of the EDG, reduce costs, and enable the loading in a fast and efficient manner. The simulation results are shown in Figs. 1.6, 1.7, 1.8, and 1.9.

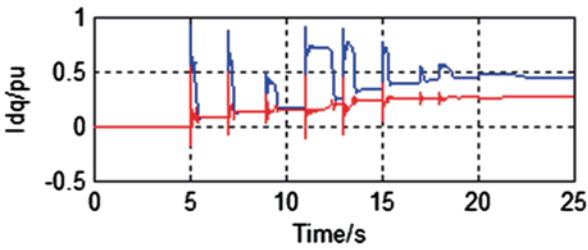


Fig. 1.7 Terminal current curve

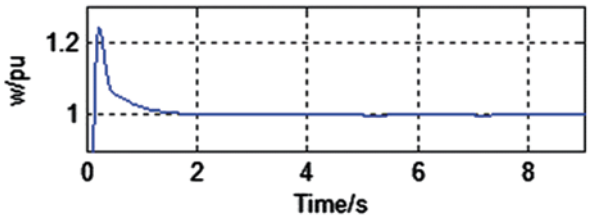


Fig. 1.8 Genetator speed curve

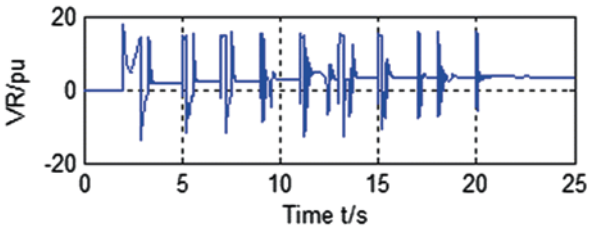


Fig. 1.9 AVR output voltage curve

The diesel engine starts first to drag the synchronous motor to synchronous speed, the exciter starts at 2 s, the first asynchronous motor is loaded at 5 s, and the other induction motors are loaded sequentially at 7, 9, 11, 13, 15, 17, 18, and 20 s.

The synchronous motor terminal voltage has the rise time of 0.93 s, the peak time of 1.11 s, the overshoot of 9%, the adjusting time of 1.37 s (deviation of 2%). The generator rotor speed has the overshoot of 24%, the rise time of 0.14 s, the peak time of 0.24 s, and the adjustment time of 1.02 s (deviation of 2%). The excitation system starts to work at 2 s, and V_R reaches its maximum after 0.1 s.

The figures show that the inrush current is large when the induction motor is loaded. As soon as the EDG loads the asynchronous motor, the stator current rapidly rises to the maximum amplitude, 6–8 times of its rated value; in addition, the effect generated by the loading process on the synchronous generator rotor speed is not significant.

We can know that the power factor of the diesel generator system becomes smaller and smaller with the loading induction motor step by step because the asynchronous motor is inductive. During the process of loading, the amplitude of voltage is stable, the adjustment of the exciter is timely and accurate, and the speed fluctuation is relatively small. At the end of the loading process, the load rate is 32% with the maximum load rate of 51%; thus, the simulation system has a two-fold margin to meet the actual demand.

1.4 Conclusion

As the EDG is an important facility to ensure the safety of nuclear power plants, it is significant to simulate and analyze the EDG. The author has established a simulation model of the EDG based on the MATLAB/Simulink platform and simulated the

loading process of the induction motor with the simulation results showing that the model is reasonable and correct; besides, the results provide a method and basis for the capacity choice of diesel engines and generators.

References

1. Ekanayake JB, Holdsworth L, Jenkins N. Comparison of 5th order and 3th order machine models for doubly fed induction generator (DFIG) wind turbines. Electr Power Compon Syst. 2003;67(2003):207–15
2. Li D, Zhang J, HE ZB. Modeling and simulation of the power system on practice teaching ship. J Shanghai Jiaotong Univ. 2008;42(02):190–3. (In Chinese).
3. Fang SL, Zhu F. Mathematical models and parameters of large turbo AC exciter. Power Syst Technol. 1986;12(04):48–58. (In Chinese).
4. Kunder PS. Power system stability and control. New York: McGraw-Hill; 1994 p. 238–9.
5. Ju P. Theory and practice of power system load modeling[J]. Autom Electr Power Syst. 1999;23(19):1–7. (In Chinese).

Chapter 2
The Metal Oxide Arresters Online Monitoring System Based on WIA-PA

Dan Li, Jie Tong, Mingyue Zhai and Yanhong Zhao

Abstract Metal oxide arresters (MOA) are widely used to protect the power facilities' overvoltage in substation, but it is very hard to find the fault point due to its large amount of usage and wide distribution. Thus, a kind of wireless networks for industrial automation-process automation (WIA-PA) online monitoring system for MOA is developed. First, the monitoring device contacts the gateway via wireless; then, the software is designed to realize the real-time monitoring of MOA situation, and finally, the test and analysis are carried out.

Keywords MOA · Online monitoring · WIA-PA

2.1 Introduction

The Metal oxide arrester's (MOA) operating quality will directly influence the security of the power system. In the operating voltage, the leakage current will flow through the arrester. In this case, the heat generated will increase the temperature of the resistor and the long-term work will make the resistors aging [1]. Once the overvoltage is produced, the arrester will be subject to the heat collapse or even an explosion, followed by the loss of the protective effect; therefore, we need uninterrupted online monitoring to ensure the safe operation of the arrester.

Back to the regular maintenance phase, generally through the nonperiodic preventive test, charging test, etc., there are big shortages represented from the following two aspects: The voltage applied to the sides of the arrester and the withstand voltage in the normal working state are different, and then, the measured data cannot accurately reflect the operating condition of the arrester; there is heavy workload and big security risk for staff. If the economic considerations and basic equipment are not enough, it will easily cause excess or lack of maintenance.

D. Li (✉) · M. Zhai · Y. Zhao
North China Electric Power University, 102206 BeiJing, China
e-mail: LiDan376@163.com

J. Tong
China Electric Power Research Institute, 100192 Beijing, China
e-mail: tongjie1@cepri.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_2

Thus, the traditional nonperiodic preventive test and charging test are unable to detect a fault timely; it is more difficult to detect the possible failures [2].

If the arrester monitoring system uses the traditional wired network as the communication mode, the staff should wire the on-site power lines and communication lines before establishment of the station by clearing the cable channel and distribution. In the course, the cabling and wiring work is more complicated, even it is heavy workload with high costs. As the long-distance transmission of analogy will introduce a lot of interferences, it is difficult to ensure the validity of the results.

The use of wireless communication is an effective solution to these problems. In this system, the industrial application of wireless networks for industrial automation-process automation (WIA-PA) technology enables wireless transmission substation in the local area with signal-synchronized capture and data-centralized transmission. We can easily achieve the system's requirements on availability; besides, it can fully comply with the unmanned substation and the demanding of state maintenance.

2.2 WIA-PA Technical Overview

WIA-PA is oriented to the information interaction among the equipments, is applied to bad industrial field environment [3]. It was put forward by Shenyang Automation Research Institution of Chinese Academy of Sciences and other 10 units; currently, it has become the official IEC standard by unanimous vote on October 14, 2011, as one of the important standards in the fields of international industrial wireless technology.

WIA-PA network of self-organization and self-healing smart mesh networks with ad hoc network routing mechanisms are very similar. As to the mesh networks in a number of communications [4], the key factor affecting network reliability is the routing. It features the property of real-time communication, high reliability, low power consumption, etc. WIA-PA network uses a double mesh and star network architecture, as shown in Fig. 2.1. It combines the centralized and the distributed management as the network resources managing methods. In this manner, the WIA-PA network has good performance and will become a focus of study in the industrial applications.

Accordingly, as demanded by the arrester online monitoring, WIA-PA can solve such problems of monitoring as the leakage current of the arrester. As a wireless communication network, each of the voltage sensor, the current sensor, and the gateway all contacts via wireless, which will not give rise to any additional risk and has a high degree of system reliability; meanwhile, according to the superiority of WIA-PA, the new online monitoring of industrial wireless communication systems can be designed so as to promote the upgrading of traditional monitoring network [5].

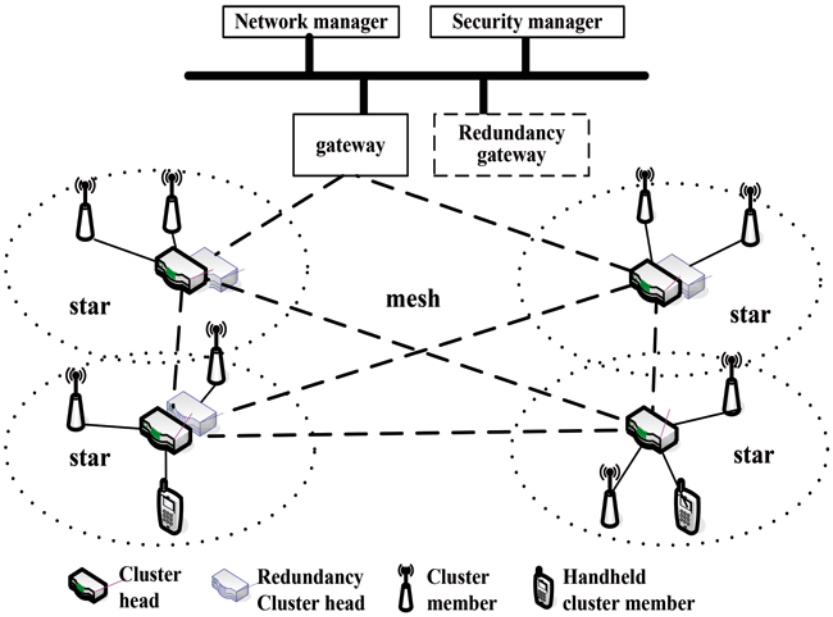


Fig. 2.1 Wireless networks for industrial automation-process automation (WIA-PA) network topology architecture

2.3 Monitoring System

This MOA online monitoring system conducts a comprehensive hardware and software design and gives the system's architecture designed principles and main module software flow diagram.

2.3.1 Structure of the Monitoring System

The online monitoring system's overall structure refers to IEC61850 protocol. It is divided into three parts: the station control layer, the spacer layer, and the process layer. And the distribution of this work is as shown below: The arrester monitoring device is mounted on each arrester's discharged valve plate and located in the process layer, which is transmitted to the wireless gateway via WIA-PA. The wireless gateway, mounted on the spacer layer, is responsible for data acquisition of the corresponding area through the short-distance WIA-PA while interacting the information with the software in PC. The software, installed on a PC which is mounted on the station control layer, can display the operational status of the arrester visually. The architecture is shown in Fig. 2.2:

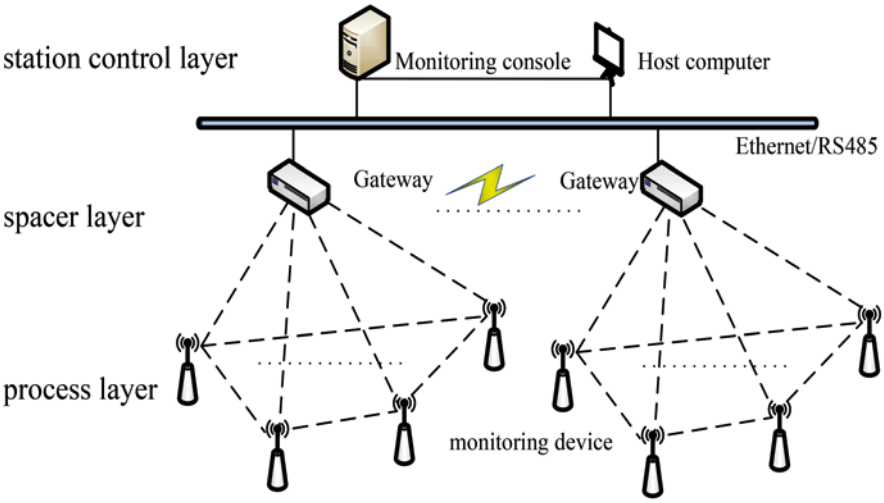


Fig. 2.2 Online monitoring system architecture

2.3.2 Arrester Monitoring Device

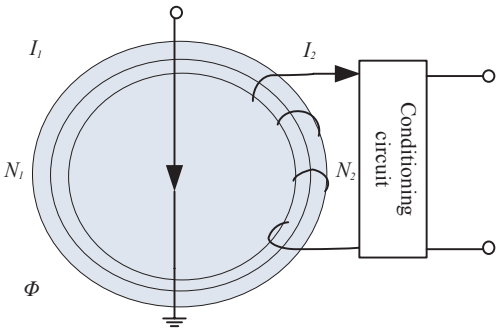
The arrester monitoring device is composed of several units [6]: current sensor, signal conditioning circuit, A/D sampling circuit, rectifier, limiting optical isolation circuit, lightning counter circuit, processor, WIA-PA communication module, display module, and solar-powered component.

The small active zero flux current sensor is connected to the ground line, extracting the current signal from the bottom of the arrester, as shown in Fig. 2.3:

The signal conditioning circuits consist of the current to voltage converter circuit, the analog filter circuit, and the enlargement circuit, as shown in Fig. 2.4.

The WIA-PA communication module operates in 250 kbps and 2.4 GHz (the free band). Its receiving sensitivity can reach −90 dBm with the signal transmitted to the WIA-PA gateway. And the solar-powered component is shown in Fig. 2.5

Fig. 2.3 Current signal extracting circuit



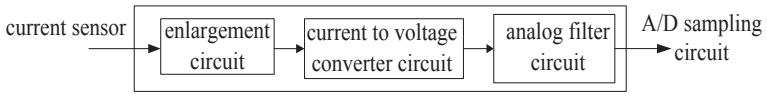


Fig. 2.4 Signal conditioning circuits

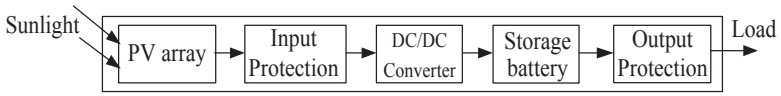


Fig. 2.5 Solar-powered component

2.3.3 WIA-PA Gateway

The furthest communication distance of intelligent WIA-PA gateway and monitoring device can reach 200 m. Each intelligent WIA-PA gateway can manage the multiple monitoring devices. Usually, at the network level of spot WIA-PA monitoring device, we use advanced and standard-based encryption, identification, authentication, key management, anti-jamming technology, etc, including the normal operation of the main wireless gateway and a redundant gateway device (to real-time backup the network information of the main gateway so that it can timely replace the original gateway and continue the network maintenance when an exception occurs).

Intelligent WIA-PA wireless gateway adopts the wired cable communication to (namely station level) PC to support Ethernet or RS485 interface way for the application of station level; then, the update cycle of data can also be configured to support the update cycle of less than 1 s. In the engineering design, compared to a wired network, the institution of the entire communication network is more concise and the construction of the network is convenient and much fast [7].

Intelligent WIA-PA wireless gateway mainly has three functions:

1. Being responsible for the management of network resources, namely, to manage the access and leave of the whole field device nodes in wireless sensor network; form and configure the WIA-PA network for the device nodes.
2. Security management of the network performance, namely, to monitor the security of the whole network.
3. Maintenance of network database, namely, to process the data in the network

2.3.4 Host Computer Software

The host computer software will firstly regard the physical addresses of the WIA-PA module or the WIA-PA gateway as the index to save the received data to the corresponding database table. According to the type of data, the software will do

the corresponding analysis and statistics and display the exceptions and warnings on the visual interface [8].

It has multiple features such as data acceptance, data storage, data display, date query, data printing, etc.

The software has the following features:

4. Data acceptance: mainly to program based on socket, be responsible for receiving packets sent by WIA-PA wireless gateway, and send control command from the monitoring center.
5. Data storage: mainly to be responsible for storing the data such as acquisition time and acquisition value in the corresponding form.
6. Data display: to curve the collected lightning arrester online monitoring parameters and show network topology and so on.
7. Date query: to generate any format of the statistics, summary tables, such as histogram by calling the database files; data query and statistics can choose any period of time; data query and statistics can define the equipment item number range by themselves.
8. Data printing: These arched data and forms can be previewed and printed, etc.

The remote request primitive is as below:

```
NLME-INFO_GET.  request(
                        DstAddr,
                        AttributeID,
                        AttributeMemID,
                        FirstValueStorIndex,
                        Count
                       )
```

The test results are as below:

2.4 Conclusion

It can be seen from the test results (Table 2.1) that the WIA-PA communication module can work best in the 2.4 GHz band and the communication interval is less than 1 s to ensure the good communication performance; under the power of solar battery, the system can work at least 2–3 years and the communication distance can be up to 800 m in the outdoor and 200 m indoor.

The WIA-PA standard is introduced in this system. With the modular design method to design an industrial wireless network system based on WIA-PA standard, the system includes most of features of WIA-PA standard. The experiment indicates that the system runs stably and further development of WIA-PA network is possible; at the same time, it verifies the design of WIA-PA standard is reasonable

Table 2.1 Test results

Communication distance (m)	RSSI	Packet loss rate (%)	Communication interval (s)
50	ec-f6	0.0	0.1
100	00-ff	0.0	0.2
200	00-ff	0.02	0.5

and realizable, and it could be widely used in various domains under WIA-PA situations, such as industrial control flow and process monitoring of industrial safety.

Acknowledgements This work was financially supported by: (1) National Development and Reform Commission ([2012]2766); and (2) Innovation Research Program of State Grid Corporation ([2012]114) and Science & Technology Program of Electric Power Research Institute (XX83-14-001).

References

1. Liang C Lina D Jianlin H Qin H Jian-guo H Yunfeng X. Development of on-line monitoring device for 10 kV MOA used in substation[J]. Insul Surge Arresters. 2012;(4):85–90 (In Chinese).
2. Hongying G Qiang G Fang S et al. The application of on-line inspection system on arrester[J]. J Electr Power. 2004;19(1):72–73 (In Chinese).
3. GB/T 26790.1-2011 Industrial wireless networks WIA specification—part 1:WIA system architecture and communication specification for process automation (WIA-PA) (In Chinese).
4. Yang L Zhijun S Peng Z, Hualiang Z. The pipeline leak monitoring system based on lndustrial wireless network WIA technology. Instr Stand Metrol. 2009;(6):31–33 (In Chinese).
5. Zhijun S Qiushi W Shijie C Peng Z. Chapter 37: on-line monitor of inteligent substation in winless communication system. Instr Stand Metrol. 2013;(1):33–34 (In Chinese).
6. Tao P, Jihong Z, Lei X. Development and application of intelligent metal oxide arrester on-line monitoring device. Manuf Autom. 2013;(10):37–39 (In Chinese).
7. Miao Y, Wei L, Weijie X. Gateway device in industrial wireless network based on WIA-PA. Comput Eng. 2010;36(23):258–261 (In Chinese).
8. Hanxin Z, Raghuveer MR. Influence of representation model and voltage harmonics on metal oxide surge arrester diagnostics. IEEE Trans Power Deliv. 2001;16(04):599–603.

Chapter 3
Two-Stage Optimal Scheduling Model of Hydrothermal Electricity Based on Generation Performance

Jiayu Li, Ruochen Li, Chao Qin, Minfang Huang, Guan Wang and Zhongfu Tan

Abstract Electric power industry is energy-intensive industry. The energy efficiency improvement, change of energy utilization mode, and reduction of carbon-based energy use of electric power industry play important roles in alleviating the contradiction between supply and demand of energy and improving the ecological environment. Theory and practice indicated that the priority use of clean and renewable energy and making cleaning power and thermal power combined operation are effective ways to make the power system economic operation, improve energy utilization efficiency, and save nonrenewable energy. At present, hydroelectric energy is in dominant among clean and renewable energies in China. Hydrothermal power system combined operation is an effective way to enhance the energy utilization efficiency of power systems and promote the adjustment of energy utilization mode. The rational coordination of hydrothermal power system operation mode which can make the best use of the complementary advantages of power systems can promote power system energy-saving operation. Tightly focusing on combined optimization dispatch problems for hydrothermal power system, this chapter has made deep research on power supply and demand as well as power output and replacements of hydrothermal units for the purpose of water resources reasonable utilization and coal and non-renewable energy conservation.

Keywords Hydrothermal units · Optimal scheduling · Energy conservation · Ideal elements extension model

C. Qin (✉) · J. Li · R. Li · M. Huang · G. Wang · Z. Tan
School of Economics and Management, North China Electric Power University, 102206 Beijing, China
e-mail: qinchao08@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_3

3.1 Introduction

In the current electricity market environment, the power generation scheduling still gives priority to the contract power mode, although some provinces promote energy-saving generation scheduling, and due to the constrained resource conditions, the applicability of energy-saving generation scheduling has been greatly restricted. So during hydrothermal generation scheduling, the optimal replacement mode still needs to be considered under the contract hydrothermal electricity generation [1]. Therefore, in this section, we construct two-phase optimized scheduling model of fire and water turbines based on generation performance. In the first stage, we discuss the relationship between company's purchases of electricity and state electricity distribution and set different power supply and demand situations and make power arrangements. In the second stage, in order to improve the overall system performance, the power trading center leads to make power replacement among turbines and then comes to the conclusion of actual electricity distribution.

Hydrothermal joint optimization scheduling in power system is under certain constraints, by optimizing certain criteria reasonably allocable share of hydropower and thermal power output, in order to achieve a certain predetermined goal [2].

Power industry implement market-oriented which adds some economic constraints in power system operation integration of new constraints, it impacts the size of the unit output, the power grid operation mode, power flow distribution, and ultimately affects the way the use of energy, efficiency, and economics of grid operation.

General hydrothermal optimization objectives are minimizing system power consumption, minimizing the cost of consumption in the entire system, the maximum profit from the sales generation side, the minimum purchase cost, the minimizing equivalent purchase cost in power plant, and other targets.

Hydrothermal power system optimization scheduling model has properties of massive, strong nonlinear. For solving them effectively, there are two types of effective optimization method [3, 4]. One is the traditional optimization algorithms, including heuristic methods, such as a slight increase rate method, Lagrange multiplier method, network flow programming, linear programming, dynamic programming, triangle cycle algorithm, triangle gyration algorithm, and coordination algorithm. The other one is to simulate the biological characteristics of some intelligent algorithms, including genetic algorithms, artificial immune algorithm, artificial neural networks, simulated annealing method, chaos algorithm, particle swarm optimization, fuzzy optimization, ant colony algorithm, differential evolution algorithm and so on [5, 6].

In this chapter, the main contents are to consider the connection in the purchase of electricity by electric company, contractual generating capacity of generation units, and the maximum generating capacity [7]. After subdividing, in the comprehensive integration of the methodology, aiming at maximizing the performance of generating units, we establish an ideal element extension model of turbine power output and displacement to analyze and calculate. According to the data of a domestic regional hydrothermal generators' operation, verify the feasibility of the model.

3.2 To Set the Situations of Electricity Supply and Demand

To set the situations of electricity supply and demand, you need to discuss the connection in electricity demand, the company's actual intended power purchase, and maximum power output, which mainly are in the following three situations.

3.2.1 Purchase of Electricity is Over than the Maximum Generating Capacity

Situation 1: In the region, the purchase of electricity by Electric Power Company outstrips the sum of maximum generating capacity of each unit. This time, in order to meet the electricity demand, we need the purchase of electricity outside the region. The regional units just need to generate in accordance with the maximum generating capacity. We do not need to consider the power generation performance in inter-unit power output. The main task is to meet the demand for electricity.

3.2.2 Purchase of Electricity is Between the Contractual Generating Capacity and the Maximum Generating Capacity

Situation 2: As the purchase of electricity by Electric Power Company below the sum of maximum generating capacity of each unit, we could divide it into two situations. One is that the purchase of electricity is over the contractual generating capacity, and the other is lower than it. In the first situation, all the units need to generate in accordance with the contractual generating capacity. The surplus is allotted in different units according to the power performance. We should organize the power output in each unit that aims at maximizing the surplus power performance.

3.3 Optimization Model of Power Output in Unit

For ease of calculation, in this section, we take a year as a scheduling circle and suppose the purchase of electricity is Q_b, the generating capacity of all the units is Q_s. While purchasing, we should consider the transmission power loss θ, and the relation of electricity supply and demand is as follows:

$$Q_b = Q_s/(1-\theta) \tag{3.1}$$

Furthermore, we suppose that there are I generators in the system. Among them, I_1 is thermal power units, I_2 is hydropower sets, $I_1 \bigcap I_2 = I$, the ith thermal power unit's maximum generating capacity is $V_i^{\max}$, annual generation hours is T_i^0, maximum hours is $T_i^{\max}$, feed-in tariff is p_{ci}, auxiliary power ratio is α_i, and the performance value is λ_i; the jth hydropower set's maximum generating capacity is $V_j^{\max}$, annual generation hours is T_j^0, maximum hours is $T_j^{\max}$, feed-in tariff is p_{hj}, auxiliary power ratio is β_j, and the performance value is φ_j. National electricity distribution and maximum generation capacity of units is calculated as follows:

$$Q_i^0 = V_i^{\max} T_i^0 (1-\alpha_i) \tag{3.2}$$

$$Q_j^0 = V_j^{\max} T_j^0 (1-\beta_j) \tag{3.3}$$

Annual maximum quantity of electricity is:

$$Q_i^0 = V_i^{\max} T_i^{\max} (1-\alpha_i) \tag{3.4}$$

$$Q_j^0 = V_j^{\max} T_j^{\max} (1-\beta_j) \tag{3.5}$$

3.3.1 Situation 1

In this situation, $Q_b \geq \sum_{i=1}^{I_1} Q_i^{\max} + \sum_{j=1}^{I_2} Q_j^{\max}$, each unit needs to generate in accordance with the maximum generating capacity. The actual quantity of on-grid electricity in each unit is:

$$Q_i = Q_i^{\max} \tag{3.6}$$

$$Q_j = Q_j^{\max} \tag{3.7}$$

3.3.2 Situation 2

In this situation, $Q_b < \sum_{i=1}^{I_1} Q_i^{\max} + \sum_{j=1}^{I_2} Q_j^{\max}$ and $Q_b \geq \sum_{i=1}^{I_1} Q_i^0 + \sum_{j=1}^{I_2} Q_j^0$, all the units need to generate in accordance with the contractual generating capacity, and the surplus is allotted in different units according to the maximum power performance. The surplus quantity of electricity is:

$$\Delta Q = Q_b - \sum_{i=1}^{I_1} Q_i^0 - \sum_{j=1}^{I_2} Q_j^0 \tag{3.8}$$

This time, the target is to maximize the performance value. The special model is:

$$\max f = \sum_{i=1}^{I_1} \lambda_i Q_i + \sum_{j=1}^{I_2} \varphi_j Q_j \tag{3.9}$$

The specific constraint conditions are as follows:

$$\sum_{i=1}^{I_1} Q_i + \sum_{j=1}^{I_2} Q_j = \Delta Q \tag{3.10}$$

$$0 \le Q_i \le \Delta Q_i \tag{3.11}$$

$$0 \le Q_j \le \Delta Q_j \tag{3.12}$$

In the formula,

ΔQ_i represents the i th thermal power unit's remaining power capacity which is over than the contractual capacity;

ΔQ_j represents the j th hydropower unit's remaining power capacity which is over than the contractual capacity.

Solving the above model, you can obtain the result of the other power output of each unit in this situation, and the actual on-grid electricity of each unit is:

$$Q_i = Q_i^0 + \Delta Q_i \tag{3.13}$$

$$Q_j = Q_j^0 + \Delta Q_j \tag{3.14}$$

3.4 The Optimized Generation Replacement Model of Units

In order to maximize the performance of the electric system, on the basis of the first stage, we should make further effects on power performance replacement. Suppose that granter is S and recipient is B, the performance value is $W_S$and W_B, among which $W_S = (w_{s1}, w_{s2}, \ldots, w_{sI})^T$ and $W_B = (w_{b1}, w_{b2}, \ldots, w_{bJ})^T$; the sold quantity of electricity of granter is Q_S=$(q_{s1}, q_{s2}, \ldots, q_{sI})$; the prices are $P_S = (p_{s1}, p_{s2}, \ldots, p_{sI})$; the sold quantity of electricity of recipient is Q_B=$(q_{b1}, q_{b2}, \ldots, q_{bJ})$; the prices are $P_S = (p_{s1}, p_{s2}, \ldots, p_{sI})$;

Then, the objective function of the model is:

$$\max \sum_{i=1}^{S} \sum_{j=1}^{B} (w_{bi} - w_{si}) Q_{ij} \tag{3.15}$$

Via the power performance replacement, we can get the result of the quantity of electricity of replacement in involved units Q_{ij}; furthermore, we can obtain the final generating capacity.

For the granter, the surplus is:

$$Q_{si}^{*} = Q_{si}^{(1)} - \sum_{j=1}^{J} Q_{ij} \tag{3.16}$$

In the formula,

$Q_{si}^{(1)}$ represents the granter's generating capacity in the first stage;
Q_{si}^{*} represents the granter's actual generating capacity after replacement;
Q_{ij} represents the quantity of electricity of replacement by the ith granted unit and recipient unit.

For the recipient, the generating capacity of jth unit after replacement is:

$$Q_{bj}^{*} = Q_{bj}^{(1)} + \sum_{i=1}^{I} Q_{ij} \tag{3.17}$$

In the formula,

$Q_{bj}^{(1)}$ represents the recipient's generating capacity in the first stage;
Q_{bj}^{*} represents the recipient's actual generating capacity after replacement.

3.5 Example Analysis

In order to make emulation on two-stage optimized scheduling model of hydrothermal units, in this section, we collected some relevant data of hydrothermal generators running in a region and discuss the different results of output scheduling in three situations. There is a sufficient storage capacity as well as relatively more hours in generating of hydropower units in this region. In order to enhance the grid connection on hydropower, we should decrease generating hours in thermal power units. In this section, we employ the ideal element extension model which is based on comprehensive integration model. Combined with the relative data (because that there are the large quantity of relative data, and our research is not focused on them. In this section, we only make out the results, not the relative data), we are going to evaluate the performance of power plants and apply the obtained data to the two stage optimized scheduling. In the second stage, we make the trade of generating replacement for the purpose of maximizing the performance. The specific generating performance and relative operating data are as follows in Table 3.1:

Furthermore, we take the regional actual on-grid quantity of electricity as the demanded one and then use the contractual distribution model and output arrangement. The statistics in on-grid energy of hydropower, the on-grid energy of thermal power, and electricity purchasing cost based on the results of the two optimizing scheduling models are in Table 3.2.

Table 3.1 Unit output arrangement in the three scenarios

	T-U1	T-U2	T-U3	H-U1	H-U2	H-U3
St1	125000	240000	300000	164000	776000	24000
St2	116750	222361	222200	64000	116850	457840
St3	125000	227400	295120	164000	776000	12480

Table 3.2 the comparative results of two kinds of electricity distribution model

	Thermal	Hydro	Cost
Contractual model	630350	616290	44640.665
Our model	461130	785510	41250.58

The proposed model can be effectively applied to reduce overall system power purchase costs. In terms of these units, due to the relatively higher performance of hydropower generating units, the allocated quantity of electricity based on the mentioned model is higher than the one based on the contractual model, which also verified the rationality of the model mentioned in this article.

3.6 Conclusion

In this article, it is divided into three kinds of power supply and demand situations, namely a maximum purchase of generating capacity, and all the units just need to generate in accordance with the maximum generating capacity. In the second situation, the purchase of electricity is below the contractual generating capacity, this time; the units reduce electricity generating capacity in accordance with the value of the power performance. We should give priority to the hydropower grid-connected power generation, when the hydro ones do not fulfill the surplus purchase of electricity; the thermal power units take the responsibility. In situation 3, purchase of electricity is between the contractual generating capacity and the maximum generating capacity. At this time, all the units generate in accordance with the contractual one; in addition, we should distribute the power output in each unit that aims at maximizing the surplus power performance. Furthermore, we analyze the generating replacement model, which is based on the above three situations. We employ the ideal element extension model which is based on comprehensive integration model to analyze and calculate. The model has been verified reasonably by the actual data in this region.

Acknowledgements This chapter is supported by the Project of Beijing municipal philosophy and social science planning (12JGC080).

References

1. Arroyo JM, Conejo AJ. Optimal response of a thermal unit to an electricity spot market. IEEE Trans Power Syst. 2000;15(3):1098–104.
2. Rodrigues RN, Finardi EC, DaSilva EL. Optimal dispatch of hydro generation plants via augmented Lagrangian. IEEE Trans Power Sys. 2005;3(7):1–6.
3. Farhat IA, El-Hawary ME. Optimization methods applied for solving the short-term hydrothermal coordination problem. Electr Power Sys Res. 2009;79(9):1308–20.
4. Rashid AHA, Nor KM. An efficient method for optimal scheduling of fixed head hydro and thermal plants. IEEE Trans Power Sys. 1991;6(2):632–6.
5. Sasikala J, Ramaswamy M Optimal gamma based fixed head hydrothermal scheduling using genetic algorithm. Exp Sys Appl. 2010;37(4):3352–57.
6. El-Hawary ME, Ravindranath KM. Hydro-thermal power flow scheduling accounting for head variations. IEEE Trans Power Sys. 1992;7(3):1232–38.
7. Lyra C, Ferreira LRM. A multi-objective approach to the short-term scheduling of a hydroelectric power system. IEEE Trans Power Sys. 1995;10(4):1750–55.

Chapter 4
Fault Features Study of Doubly Fed Induction Generators on the Stator Windings Turn-to-Turn Short Circuit

Junqing Li, Dong Wang, Ximei Wang and Lipeng Zhang

Abstract Doubly fed induction generators (DFIG) are gradually used in wind power generation, and its capacity is becoming bigger and bigger. Turn-to-turn short circuit of winding is common fault in DFIG, and in order to study the steady electromagnetic characteristic of DFIG with the stator windings turn-to-turn short circuit, a simulation model is built based on the finite element method. In order to improve the accuracy, the excitation current in finite element is calculated by multicircuit equation. The proposed model is employed to simulate the magnetic field and current of stator and rotor windings when turn-to-turn short circuits occur on stator windings of DFIG. From the simulated result, we obtained the frequency spectrum of magnetic flux density of air gap, analyzed the difference of magnetic field and line current of stator and rotor windings between faulty and healthy generators, and concluded the feature of stator inter-turn fault in DFIG.

Keywords DFIG · Fault feature · Finite element method (FEM) · Multi-circuit equation · Stator turn-to-turn short circuit

J. Li (✉) · D. Wang · X. Wang · L. Zhang
School of Electrical and Electronic Engineering, North China Electric Power University, 071003 Baoding, China
e-mail: junqing03@163.com

D. Wang
e-mail: 804910637@qq.com

X. Wang
e-mail: 497026977@qq.com

L. Zhang
e-mail: dapengniaoroc@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_4

4.1 Introduction

Wind power generation has been gradually increasing in the world; doubly fed induction generators (DFIG) are used in wind power generation and with capacity being also bigger and bigger, so security and reliability of DFIG become more important. Turn-to-turn short circuit of winding is one of the common faults in electric machines, so the detection of stator inter-turn short fault has drawn a lot of attention [1–4]. It summarized and compared the stator inter-turn fault detection approaches [2]. Nonintrusive condition-monitoring methods, which monitor the motor's condition using only the currents and voltages of the motor, are preferred due to their low cost and nonintrusiveness [2]. Different methods were used to study the variation of physical quantities for inter-turn fault machine [5–11]. Inter-turn short circuit of stator winding in permanent magnet motors was studied by the finite element method [5]. Finite element model was used to study inter-turn short circuits in induction motors. The wave of stator current, magnetic flux density in air gap, and torque was gave, analyzing frequency spectrum of stator current [6, 7]. Multicircuit model was employed to study stator inter-turn short circuits in DFIG with star connection stator winding, which indicated that stator inter-turn short circuit causes triplex harmonic component occurring in stator line current [8]. Multicircuit model of DFIG was presented with delta connection stator winding and analyzed current with inter-turn fault [9]. Harmonic component of rotor current was analyzed when inter-turn short circuits occur at stator winding in DFIG and used search-coil voltage signature to detect stator inter-turn fault in DFIG [10]. Time-stepping finite element method (FEM) of PM motor presented with inter-turn fault and studied magnetic field and identification of the machine parameters [11].

In the field of studying monitor, diagnosis, and protection approach when inter-turn short-circuit fault occurs in electrical machines, it is the basis to obtain the feature of the electric and magnetic quantities in machines with turn-to-turn short circuit. This chapter presents an accurate model of DFIG with delta connection stator winding, simulates steady process of magnetic field and current of stator and rotor windings, and summarizes the electromagnetic characteristic of DFIG with stator inter-turn fault.

4.1.1 Mathematical Model

To accurately have the model of the complex geometry structure and magnetic circuit saturation of DFIG, two-dimensional (2D) finite element method was employed [5–6], and the governing equation is given by (4.1)

$$\begin{cases} \Omega : \dfrac{\partial}{\partial x}\left(\nu \dfrac{\partial A_Z}{\partial x}\right) + \dfrac{\partial}{\partial y}\left(\nu \dfrac{\partial A_Z}{\partial y}\right) = -J_Z + \sigma \dfrac{dA}{dt} \\ \tau : A_Z = A_{Z0} \end{cases} \tag{4.1}$$

where A_Z and J_Z are the axial components of the magnetic vector potential and the source current density, respectively; ν is the material reluctivity; σ is the material conductivity; and A_{Z0} is the magnetic vector potential in the first boundary.

For efficient and precise determinations of the magnetic field of DFIG on the stator windings turn-to-turn short circuit, source current density of stator and rotor windings in (1) is needed, and the current of stator and rotor windings is computed by multicircuit model of DFIG. The multicircuit equations of DFIG are given by (4.2) and (4.3) [9].

$$\frac{dI'}{dt} = AI' + B \tag{4.2}$$

$$I = H^T I' \tag{4.3}$$

In (4.2), $A = (-HLH^T)^{-1}[p(HLH^T) + HRH^T]$ and $B = (HLH^T)^{-1}HU$,

where H is the conversion matrix from branch to loop, I' and I are, respectively, loop and branch currents' matrix, L and R are, respectively, inductance and resistance matrices, and U is the branch voltage matrix in stator and rotor windings.

The proposed model is utilized to study electromagnetic field and current of stator and rotor windings in DFIG, the simulated machine is 5.5 kW and three-phase generator, there are 37 turns in a single stator coil, and other parameters of generator are listed in Table 4.1. The study concentrates on the turn-to-turn short circuit occurring in a single coil because it is the incipient stage of severe fault, which is the basis for studying other severe inter-turn faults. When inter-turn fault occurs at a branch of stator phase winding (e.g., C phase), the number of loop and branch in multicircuit model increases one than that of healthy generator, which is shown in Fig. 4.1.

Table 4.1 Parameters of the simulated generator

Parameter	Value	Unit
Rated power	5.5	kW
Rated frequency	50	Hz
Rated voltage	380	V
Number of pole pairs	2	
Stator slots	36	
Stator winding	△ connection	
Stator parallel branchs	2	
Rotor slots	24	
Rotor winding	Y connection	
Rotor parallel branchs	1	

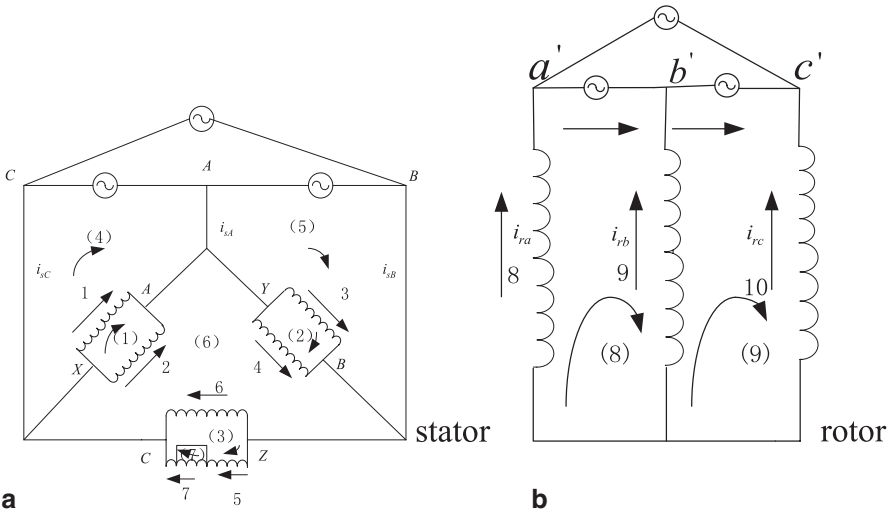


Fig. 4.1 Choice about branch and loop of generator with inter-turn fault

4.2 Simulation Results and Analysis

We studied the turn-to-turn short-circuit fault occurring in a single stator coil of stator C phase winding. The generator is modeled supposing that the rotor speed is constant when stator inter-turn short circuit occurs, a three-phase symmetrical slip frequency source is used for energizing the rotor winding, and stator winding is connected to a balanced grid. To make a clearly show of the features of the electric and magnetic under different operating conditions, the following different cases are considered in this chapter.

1. Case A: Healthy DFIG with symmetrical stator and rotor windings, rotor runs at 1560 and 1800 rpm, respectively;
2. Case B: Faulty DFIG with symmetrical rotor winding, inter-turn short circuit occurs in a single stator C phase coil, subjected to 1-turn fault, 5-turns fault, and 10-turns fault, respectively, at 1560 rpm rotor speed;
3. Case C: Faulty DFIG with symmetrical rotor winding, inter-turn short circuit occurs in a single stator C phase coil, subjected to 1-turn fault and 5-turns fault, respectively, at 1800 rpm rotor speed.

Figures 4.2 and 4.3 describe the magnetic field and magnetic flux density in the air gap for healthy machine. As shown in the graph, magnetic field and magnetic flux density in air gap are symmetrical.

As shown in Fig. 4.4, the air gap magnetic field contains 250 and 350 Hz harmonic, and very few 150 and 450 Hz harmonic. In addition to this, it also includes 550 and 650 Hz of rotor tooth harmonic, and 850 and 950 Hz of stator tooth harmonic because the developed model is 2D and rotor slot skew is neglected causing

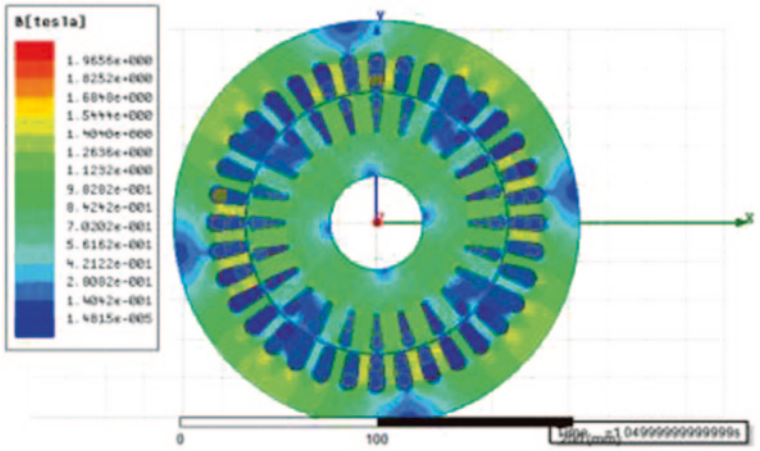


Fig. 4.2 Magnetic field for healthy generator

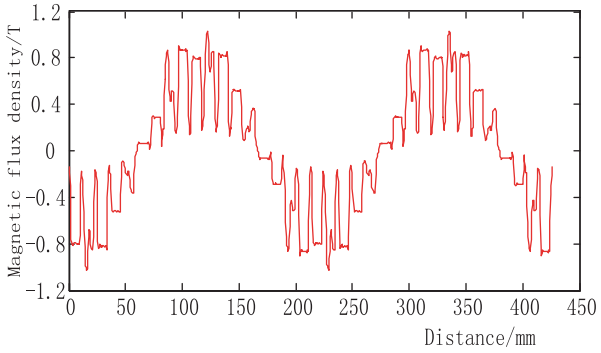


Fig. 4.3 Magnetic flux density in air gap for healthy generator

a stronger variation in the reluctance path. For healthy generator, the magnetic flux density of air gap only contains odd harmonic, not even and fractional one. Stator and rotor line currents are also symmetrical, respectively, in case A. In case B, the magnetic field, magnetic flux density in the air gap, and current of stator are asymmetrical. Due to space limitations, Figs. 4.5–4.7 give only magnetic field, magnetic flux density in the air gap, and its frequency spectrum under 5-turns shorted in a single stator C phase coil at 1560 rpm rotor velocity.

In Figs. 4.3 and 4.6, the horizontal ordinate represents space distance of air gap in unit millimeter and the vertical ordinate represents magnetic flux density in unit Tesla. From Fig. 4.2 and Fig. 4.5, comparing with the healthy generator, magnetic field is asymmetrical for the fault one, and the saturation degree in position of inter-turn fault increases.

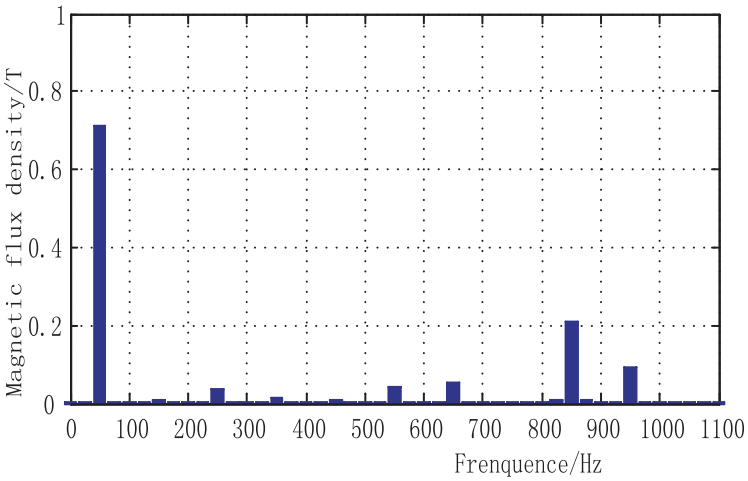


Fig. 4.4 Frequency spectrum of magnetic flux density in air gap for healthy generator

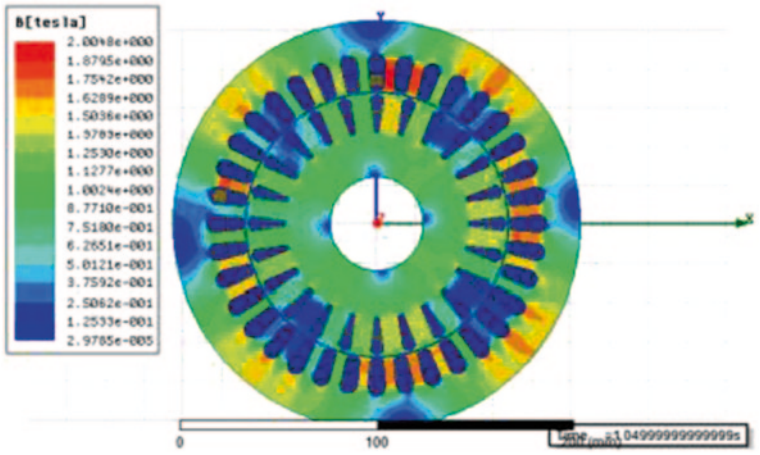


Fig. 4.5 Magnetic field for faulty generator on 5 turns shorted, in case B

From Fig. 4.7, harmonic component in faulty generator is obviously richer than healthy machine. For inter-turn fault machine, the magnetic field contains all odd, even, and fractional harmonic. In the fault machine, comparing with healthy machine, 550 and 650 Hz of rotor tooth harmonic slightly decreases, 50 Hz fundamental component, 850 and 950 Hz of stator tooth harmonic slightly increases, 150 Hz harmonic obviously increases. In addition to that, all fractional harmonics occur on magnetic, in which 75, 175, 275 Hz harmonics and so on obviously increase. Hundred hertz even harmonic and so on occur on magnetic field too.

Figures 4.8 and 4.9 describe stator and rotor line currents, stator line currents are asymmetrical and rotor line currents are symmetrical.

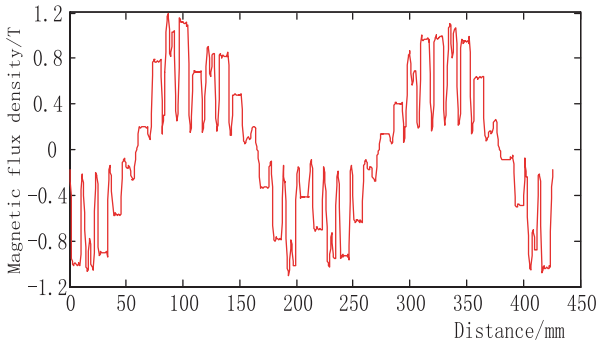


Fig. 4.6 Magnetic flux density in air gap for faulty generator on 5-turns shorted, in case B

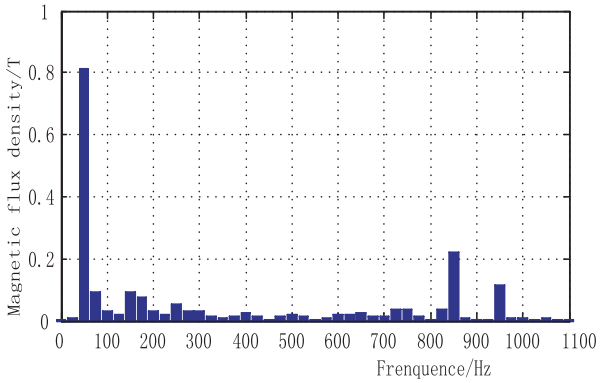


Fig. 4.7 Frequency spectrum of magnetic flux density in air gap on 5-turns shorted, in case B

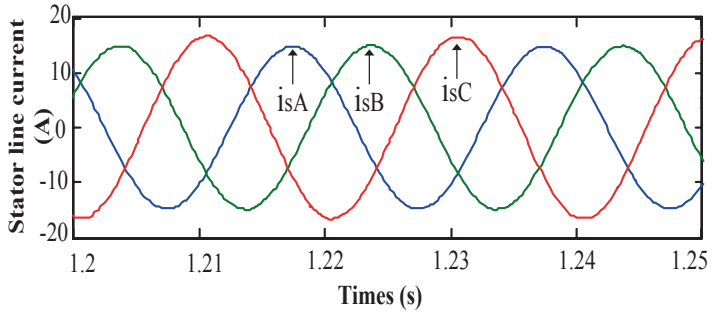


Fig. 4.8 Line current of stator windings on 5-turns shorted, in case B

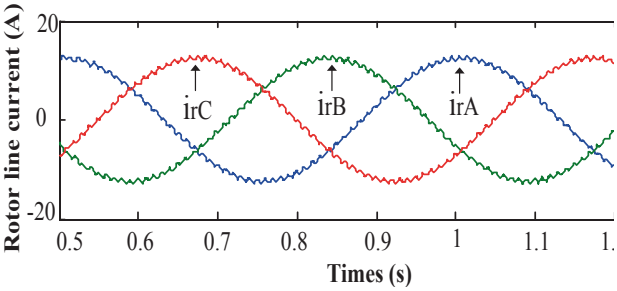


Fig.4.9 Line current of rotor windings on 5-turns shorted, in case B

Tables 4.2 and 4.3 demonstrate the difference of amplitude and phase angle of the stator line currents between healthy and faulty generators in case B. $I_{H\mathrm{m}}$ represents the line current amplitude of healthy generator, and $\frac{\Delta I_{\mathrm{m}}}{I_{H\mathrm{m}}}$ expresses the rate of change of line current amplitude in faulty generator relative to the healthy one.

As deduced from Table 4.2, it is clear that stator line current increases with growing number of inter-turn shorted and stator C line current I_{Cm} is more greater than other line currents $I_{A\mathrm{m}}$ and I_{Bm} when inter-turn short circuit occurs in a branch of C phase winding. From Table 4.3, phase angle difference θ_{AB} between stator line currents i_{sA} and i_{sB} decreases and less than 120°, and the other phase angle difference θ_{BC} and θ_{CA} increases and greater than 120° with growing number of inter-turn shorted, when inter-turn shorted occurs in a branch of stator C phase winding. We simulate the current when inter-turn shorted occurs at a different phase coil, and the result indicates that the phase angle difference between the fault phase and the leading phase is greater than the others.

In case C, magnetic field, magnetic flux density in the air gap, and the current of stator and rotor windings follow the same principles with case B. Tables 4.4 and 4.5 demonstrate the difference of amplitude and phase angle of stator line currents between healthy and faulty generators in case C. The rate of change of stator line current I_{Cm} in case C is larger than in case B. The phase angle difference between the faulty phase and the leading phase is greater than that in case B on the same fault degree.

Table 4.2 Amplitude value and rate of change of stator line currents in case B

Condition	$I_{A\mathrm{m}}$ (A)	I_{Bm} (A)	I_{Cm} (A)	$\frac{\Delta I_{A\mathrm{m}}}{I_{H\mathrm{m}}}$ (%)	$\frac{\Delta I_{\mathrm{Bm}}}{I_{H\mathrm{m}}}$ (%)	$\frac{\Delta I_{\mathrm{Cm}}}{I_{H\mathrm{m}}}$ (%)
Healthy	14.1	14.1	14.1	0	0	0
1-turn shorted	14.4	14.4	14.6	2	2	3
5-turns shorted	14.8	15.0	16.5	5	6	17
10-turns shorted	14.8	15.7	17.3	5	11	22

Table 4.3 Angle difference between stator line currents in case B

Condition	θ_{AB}	θ_{BC}	θ_{CA}
Healthy	120	120	120
1-turn shorted	118.8	121.1	120.1
5-turns shorted	112.4	124.2	123.4
10-turns shorted	110.5	128.3	121.2

Table 4.4 Amplitude value and rate of change of stator line currents in case C

Condition	I_{Am} (A)	I_{Bm} (A)	I_{Cm} (A)	$\frac{\Delta I_{Am}}{I_{Hm}}$ (%)	$\frac{\Delta I_{Bm}}{I_{Hm}}$ (%)	$\frac{\Delta I_{Cm}}{I_{Hm}}$ (%)
Healthy	16.5	16.5	16.5	0	0	0
1-turn shorted	16.8	16.8	17.4	1.7	1.7	5.1
5-turns shorted	16.8	18.1	19.0	1.7	9.4	14.5

Table 4.5 Angle difference between stator line currents in case C

Condition	θ_{AB}	θ_{BC}	θ_{CA}
Healthy	120	120	120
1-turn shorted	119.6	123.4	116.6
5-turns shorted	114.7	126.6	118.7

In Tables 4.3 and 4.5, the difference of phase angle of stator line currents is defined by comparing zero-crossing point of every current wave. From the simulated result, it is deduced that the inter-turn short circuit can cause asymmetry of current in stator winding. An asymmetrical fundamental current generates reversal magnetic field, which results in harmonic occurring in magnetic field of air gap and stator and rotor currents. Amplitude of current and phase angle difference increases with growing rotor slip.

4.3 Conclusion

The proposed model was used to simulate the electromagnetic characteristic when inter-turn fault occurs in a single stator coil. This chapter analyzed the magnetic field distribution and saturation degree, magnetic flux density of air gap, and amplitude and phase angle difference of stator and rotor currents and gave the difference of these quantities between healthy and faulty generators. The simulated result indicated that:

1. Inter-turn fault of stator winding causes faulty branch current to substantially increase.
2. The phase angle difference between the fault phase and the leading phase is greater than the others, over 120°, which increases with growing number of inter-turn short circuits.
3. Stator inter-turn fault results in the variation of harmonic component in air gap magnetic field. This chapter summarized characteristic of all harmonic components.

Acknowledgements This work was supported by Natural Science Foundation of Hebei Province in China (E2014502015).

References

1. Ribrant J, Bertling LM. Survey of failures in wind power systems with focus on Swedish wind power plants during1997–2005. IEEE Trans Energy Convers. 2007;22(1):167–73.
2. Zhang P, Du Y, Habetler TG, Lu B. A survey of condition monitoring and protection methods for medium-voltage induction motors. IEEE Trans Ind Appl. 2011;47(1):34–46.
3. Gandhi A, Corrigan T, Parsa L. Recent advances in modeling and online detection of stator inter-turn faults in electrical motors. IEEE Trans Ind Electron. 2011;58(5):1564–75.
4. Nemec M, Drobnic K, Nedeljkovic D, Fiser R, Ambrozic V. Detection of broken bars in induction motor through the analysis of supply voltage modulation. IEEE Trans Ind Electron. 2010;57(8):2879–88.
5. Babak V. Fault analysis and parameters identification of permanent-magnet motors by the finite element method. IEEE Trans Magn. 2009;45(9):3290–95.
6. Rojas C, Melero MG, Cabanas MF, Cano JM, Orcajo GA, Pedrayes F. Finite element model for the study of inter-turn short circuits in induction motors. IEEE International Symposium on diagnostics for Electric Machines, Power Electronics and Drives, IEEE Conference Publications, Cracow; 2007. pp. 415–9.
7. Abdesselam L, Ammar M, Guy C. Analysis of inter-turn short circuit in slots by finite element model. 2011 10th International Conference on Environment and Electrical Engineering (EEEIC), IEEE Conference Publications Rome; 2011. pp. 1–4.
8. Lu QF, Cao ZT, Ritchie E. Model of stator inter-turn short circuit fault in doubly-fed induction generators for wind turbine. IEEE 2004 35th Annual IEEE Power Electronics Specialists Conference, IEEE Conference Publications, Aachen Germany; 2004. pp. 932–7.
9. Li J, Wang D, Ma L. The research of the inter-turn short circuit of the stator windings in doubly fed induction generator. 2011 International Conference on Electrical Machines and Systems(ICEMS 2011), IEEE Conference Publications, Beijing, China; 2011. pp. 1–4.
10. Dhaval S. Stator-inter-turn-fault detection of doubly fed induction generators using rotor-current and search-coil-voltage signature analysis. Conference Record of the 2008 IEEE Symp.0n.Electr. Insulation. IEEE Conference Publications, New Orleans, LA; 2007. pp. 47–51.
11. Vaseghi B, Nahidmobarakeh B, Takorabet N, Meibody-Tabar F. Inductance identification of PM motor with winding turn short circuit fault. 2010 14th Biennial IEEE Conference on Electromagnetic Field Computation (CEFC), IEEE Conference Publications, Chicago, IL; 2010. pp. 1–1

Chapter 5
Estimation of the State of Charge of the Battery Based on Driving Cycles Discriminant

Niaona Zhang and Zhe Zhang

Abstract It is the key technical parameter for the battery management system in electric vehicles to estimate the state of charge (SOC) of batteries. It is difficult to establish an accurate mathematical model due to the influence of characteristic of monomer battery, consistency of batteries, and balance control technology. First, the driving cycles of the vehicle are classified by the Bayes classification method; secondly, the SOC prediction model of multi-scale support vector machine based on the driving cycle discrimination is constructed. According to the statistical characteristics of different driving cycles, the model parameters are optimized by Levenberg–Marquardt algorithm to improve the prediction accuracy of SOC. Finally, the rationality and practicability of the proposed method are verified through simulation and analysis.

Keywords Driving cycles · SOC estimation · Support vector machines · Levenberg–Marquardt

5.1 Introduction

The storage battery is one of the key technologies for electric vehicles. How to improve the accuracy of state of charge (SOC) estimation is the key technique for the battery management system and it is related to many factors: the functions of SOC including the density of the battery electrolyte, the open-circuit voltage, the resistance, and other parameters. In order to improve the accuracy of SOC estimation, many kinds of methods have been proposed such as the discharge test method [1], Ah metrology, the open-circuit voltage [2], the load voltage method [3], the electrochemical impedance spectroscopy [4], the resistance method, the linear model approach, the neural network and so on [5]. The traditional method of battery

N. Zhang (✉) · Z. Zhang
Changchun University of Technology, 130012 Changchun, China
e-mail: zhangniaona@mail.ccut.edu.cn

Z. Zhang
e-mail: zhangzhekk@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_5

estimated performance is based on precise mathematical models, but the establishment of it is based on the physical and chemical regulations reacted in the battery internal. It applies a large number of the assumptions and empirical parameters to the process of modeling, and the model precision is limited. And the expression of the model is a multi-parameter set of partial differential equations, and the solving process is so complicated to solve some simpler problems. As to the battery under the actual working state, the internal reaction is very complex and it is very difficult to meet the actual needs.

Support vector machines (SVM), as developed in the statistical learning theory, have been widely used in terms of nonlinear classification, function approximation, and pattern recognition [6, 7]. Compared with the neural network, the SVM method has more solid theoretical foundation of mathematics, and it can effectively solve the problem of high-dimensional data model constructed under limited conditions; besides, SVM also features the advantage of generalization ability, convergence to the global optimum, and the dimension of insensitivity.

Levenberg–Marquardt (LM) algorithm, as a kind of iterative technique to locate the minimum variable function, is one of the basic technologies of solving nonlinear least squares problems. It can deal with redundant parameters effectively, and it has been widely used in the optimization algorithm.

In this paper, in combination with characteristics of the storage battery remaining energy in HEV which is used as bus in one city, the prediction mode is established based on the least squares support vector machine. According to different driving cycles, the established prediction model is decomposed under multi-scale. As to different battery discharge modes, a different-scale kernel function is adopted to implement the prediction of the battery remaining capacity with its structure as shown in Fig. 5.1.

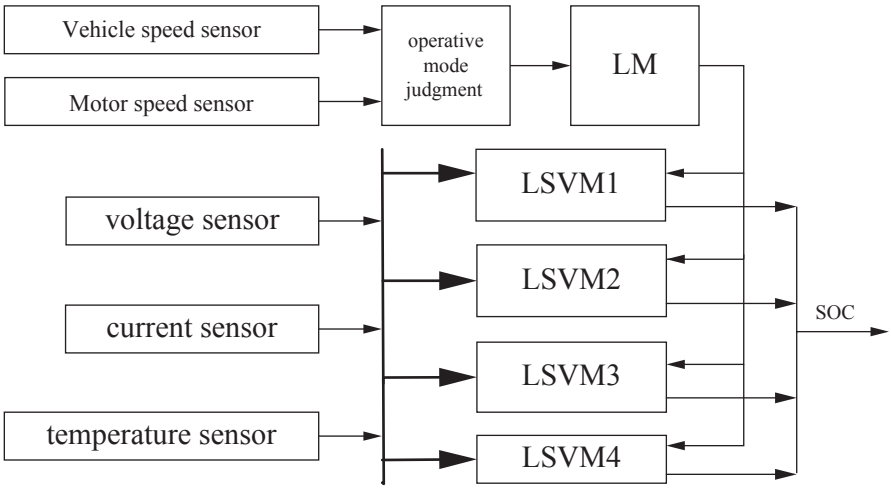


Fig. 5.1 Based on least squares support vector machine of the battery remaining capacity estimated

5.2 Prediction Model of Battery Remaining Capacity Based on Least Squares Support Vector Machine

The idea of SVM for multivariate nonlinear estimation is to introduce the hyperplane thought to map all of the nonlinear estimation problems into a hyperplane linear estimation. The input of the system is the ambient temperature, battery voltage, and current $x = [x_1, x_2, x_3]$, while the output of the system is the battery remaining capacity $y = [y]$. Based on the least squares SVM, the steps of the prediction of battery remaining capacity are shown as below:

The kernel function is selected as the quadratic polynomial:

$$k(x_i) = (1 + x^T x_i)^2 \tag{5.1}$$

where the category of data that are acquired by the input sensor is 3, $x_i = [x_{i1}, x_{i2}, x_{i3}]$, and n is the number of data samples, $i = 1, 2, \ldots, n$.

The Lagrange function is defined as:

$$G = \frac{1}{2}\omega^T \omega + \frac{C}{2}\sum_{i=1}^{n} e_i^2 - \sum_{i=1}^{n} \alpha_i [y_i - \omega^T k(x_i) - b - e_i] \tag{5.2}$$

where ω is the weight vector, b is the bias, and ω and b are the undetermined parameters. e_i is the accuracy of the training data, and e_i^2 is the least squares loss function. C is the capacity factor. α_i is the Lagrange multiplier, and x_i and y_i are the known input–output data samples, where $y_i = [y_i]$. In order to reach the maximum of G, used $\frac{\partial G}{\partial \alpha_i} = 0$, $\frac{\partial G}{\partial b} = 0$, $\frac{\partial G}{\partial e_i} = 0$, $\frac{\partial G}{\partial \omega} = 0$ to simultaneous equation:

$$\frac{\partial G}{\partial \omega} = 0 \Rightarrow \omega = \sum_{i=1}^{n} \alpha_i k(x_i) \qquad \frac{\partial G}{\partial b} = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i = 0 \tag{5.3}$$

$$\frac{\partial G}{\partial e_i} = 0 \Rightarrow \alpha_i = C e_i \qquad \frac{\partial G}{\partial \alpha_i} = 0 \Rightarrow \omega^T k(x_i) + b + e_i - y_i = 0$$

where $i = 1, \ldots, n$. Remove the parameters ω and e_i, and the equation becomes

$$\begin{bmatrix} 0 & L^T \\ L & K + C^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{5.4}$$

where $L = [1 \ldots 1]^T$, $Y = [y]^T$, $\alpha = [\alpha_1 \ldots \alpha_n]^T$, $K = k^T(x_i)k(x_j)$ is a matrix. $j = 1, 2, \ldots, n$. According to the Formula (5.4), we will obtain the parameters α and b with the method of the squares method.

According to requirements of α and b, it gets the forecast output of SOC:

$$y(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b \tag{5.5}$$

5.3 Based on the Levenberg–Marquardt Algorithm of the Multi-Resolution Support Vector Machines

A multi-scale compact expression is given through the multi-scaling decomposition of the kernel function in the objective function. It ensures the overall trend of target curve without distortion and possesses good local approximation ability as well as small generalization error. In this chapter, according to different driving cycles, optimize the best decomposition scale by the Levenberg–Marquardt algorithm to obtain the optimal scale.

The scaling kernel function is defined as:

$$K(x, x_n) = \sum_{\delta} 2^{-j} k(2^{-j} x - \delta) k(2^{-j} x_n - \delta) \tag{5.6}$$

where δ is the shift coefficient and 2^j is the scale.

Define $q = [k, 2^j]$, q^{new} is the parameter after optimization, and q^{old} is the parameter before optimization. e $= q^{new}$-q^{old} is an error. The e can be expanded into a first-order Taylor series when the q^{new}-q^{old} is very small.

$$e(q^{new}) = e(q^{old}) + \mu(q^{new} - q^{old}) \tag{5.7}$$

where

$$(\mu)_{pi} = \frac{\partial e^p}{\partial q_i} \tag{5.8}$$

where p is the pth samples and e is a vector for e^p element.

The sum of the squared error is as below:

$$E = \frac{1}{2} \left\| e(q^{old}) + \mu(q^{new} - q^{old}) \right\|^2 + \lambda \left\| q^{new} - q^{old} \right\|^2 \tag{5.9}$$

where $\lambda = 10^{\mathrm{sgn}(e(q^{new}) - e(q^{old}))}$, and it chooses the optional initial values. In order to obtain the smallest E, we take the derivative of q^{new}.

$$q^{new} = q^{old} - (\mu^T \mu + \lambda I)^{-1} \mu^T e(q^{old}) \tag{5.10}$$

Take Formula (5.10) into Formula (5.1–5.7)~(5.1–5.9) and repeat it like this until E meets the requirement; the calculated data q^{new} is the best decomposition scale k, 2^j of Formula (5.6) by using Levenberg–Marquardt algorithm at the moment.

5.4 Prediction of the Driving Cycle

The Bayes discriminant method often describes this knowledge by the prior probability distribution and draws a sample. Revise the existing knowledge by using sample to get a posterior probability distribution. Finally, the driving cycles are classified and discriminated by the posterior probability method to make the effect to be more ideal and the application more extensive. In this paper, we select the maximum speed, average velocity, minimum acceleration, average acceleration, uniform time proportion, idle time proportion, average acceleration of acceleration section and average deceleration of retardation section, the average opening of the pedal, and the time of pedal to zero to be characteristic values. The urban driving cycles are divided into the crowded and the unimpeded. The characteristic parameter in different driving cycles is acquired. Then, the driving cycles are analyzed by the Bayes classification method. First, the driving process is divided into several pieces according to certain period, and the characteristic parameters that have a greater influence on the condition prediction of each driving segment are extracted. Several typical cycles that can represent the bus driving process are divided. Then, the real-time measured data are forecasted based on the Bayes discriminant method [8].

Assume that there is the overall $G_i(i=1,2,\cdots,k)$ and it has the probability density function $f_i(x)$. According to the analysis of mathematical statistics, we know that the frequency of occurrence of G_i is q_i. That is to say, when the sample x_0 occurs, calculate x_0 belonging to what kind of probability. Calculate the posterior probability based on Bayes formula:

$$P(G_i \mid x_0) = \frac{q_i f_i(x_0)}{\Sigma q_j f_j(x_0)} \tag{5.11}$$

Discrimination rules are shown as below:

$$P(G_l \mid x_0) = \frac{q_l f_l(x_0)}{\Sigma q_j f_j(x_0)} = \max_{1\le i\le k} \frac{q_i f_i(x_0)}{\Sigma q_j f_j(x_0)} \tag{5.12}$$

If we consider the misjudgment loss, use the Bayes discriminant analysis of the misjudgment loss. If there is the overall G_i $(i=1,2,\ldots,k)$, the prior probability is $\{q_i,\ i=1,2,\ldots,k\}$, and the G_i has the probability density function $f_i(x)$; when the loss is $\{C(j/i)\}$, the solution of the divided Bayesian is as follows:

$$D_i = \left\{\mathbf{x} \mid h_i(\mathbf{x}) = \min_{1\le j\le k} h_j(\mathbf{x})\right\}, \qquad i = 1, 2, 3, \cdots, k \tag{5.13}$$

where $h_j(x) = \sum_{i=1}^{k} q_i\ C(j/i)\ f_i(x)$, and x is the sample extraction value of the unknown population. Firstly, calculate *K, the* wrong judgment average values according to the prior probability weighting, and then compare the size of them. Finally, select the smallest to determine whether this sample is general.

5.5 Simulation Analysis

Firstly, the driving cycle is divided into urban, suburban, and high speed. According to the Bayes discriminant method, we use the data sample to build a sample library about the kinematic characteristic values. Assume that the current temperature is 25°C and the initial value of SOC is 0.4. With NEDC driving cycle, the output voltage, the current, and predicted curves of soc are shown in Figs. 5.2, 5.3, 5.4, 5.5.

From the simulation diagrams, under the NEDC condition, when the driving speed reaches the maximum, the storage battery is in the state of discharge; at the same time, the value of SOC reaches the minimum. Compared with the traditional method, the predictive efficiency is improved and the predictive result is more reasonable and practical.

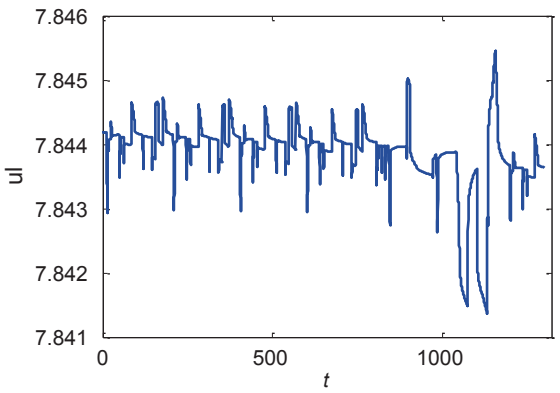


Fig. 5.2 Output voltage of the battery

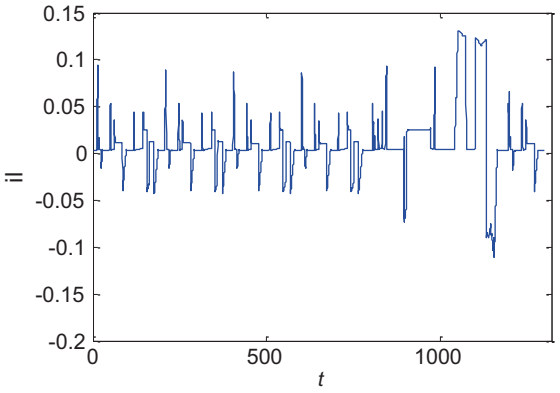


Fig. 5.3 Output current of the battery

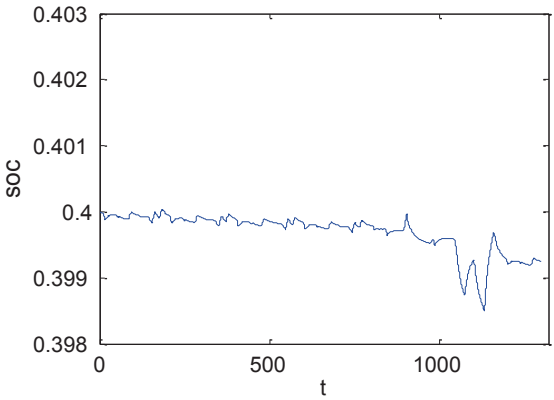


Fig. 5.4 State of charge (SOC) of battery estimate

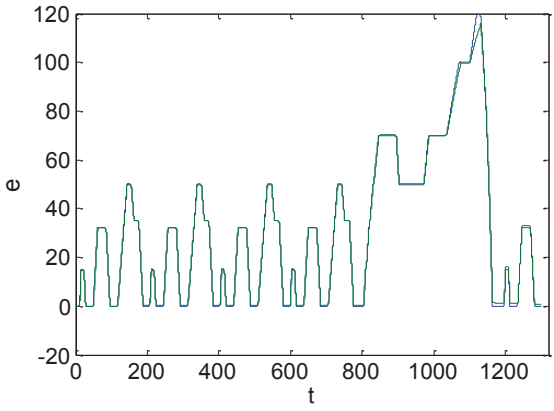


Fig. 5.5 Vehicle speed and given speed under the NEDC driving cycle

5.6 Conclusion

The prediction method of the battery remaining capacity based on the multi-scale least squares support vector machine features a higher accuracy rate. It is applicable to the classification problem in the small sample and may improve the calculation speed effectively. The data in different cycles are optimized by the improved LM algorithm. The prediction of the residual capacity in any state of nickel–metal hydride battery's charging and discharging process can be realized and it makes the forecasting result more accurate, reasonable, and practical.

References

1. Jiang P. Investigation of the driving cycle construction of mixed roads in the city. Hefei: Hefei University of technology; 2011.

2. Liu Q. Study on battery management system of MH/NI batteries and estimation of SOC. Wuhan: Wuhan University of technology; 2004.
3. Luo YT, Hu HL. Analysis and distinguish the driving cycle of HEV. J South China Univ Technol (Natural Science Edition). 2007;35(6):8–13 (In Chinese).
4. Lu Y, Fang J. Research on the model of SOC for Ni-MH battery used in electric vehicle. Chin Battery Ind. 2006;11(5):65–9 (In Chinese).
5. Li HC, Tian GY. Character of MH/Ni battery used in EV. Battery Bimon. 2002;32(5):11–5(In Chinese).
6. Zhang JM, Zeng J. Design of battery monitor system of electric vehicle. Chin J Sci Instrum. 2006;27(6):223–5 (In Chinese).
7. Zhang XL, Song JJ. RBF neural networks based on dynamic nearest neighbor-clustering algorithm and its application in prediction of MH-Ni battery capacity. Trans China Electrotech Soc. 2005;20(11):84–6 (In Chinese).
8. Deng C, Shi PF. Prediction of residual capacity of MH/Ni batteries based on neural network. J Harbin Inst Technol. 2003;35(11):1406–8 (In Chinese).

Chapter 6
Rotor Fault Slot Positioning of Doubly Fed Induction Generators Based on Wavelet Decomposition

Junqing Li, Lipeng Zhang and Haibo Yu

Abstract Doubly fed induction generator (DFIG) is widely used in the field of wind power, and the rotor windings inter-turn short-circuit fault of DFIG is one of the focuses of the current domestic and international research. Taking into account the special characteristics of the structure and operational, it is difficult to make accurate positioning of inter-turn short-circuit fault of DFIG. In this chapter, the finite element theory and Ansoft are used to simulate and analyze the diffidence of flux density between the DFIG running in normal and rotor windings inter-turn fault, getting the variation characteristics. It is indicated the feasibility of locating a fault slot from the perspective of the magnetic properties. Then, the position of fault slots is located by the data, collecting by the combination of the detection coil and optoelectronic devices, and detecting by wavelet decomposition.

Keywords Doubly fed induction generator · Inter-turn short-circuit fault of rotor windings · Wavelet decomposition · Locate the fault slots

6.1 Introduction

As the core of wind power, doubly fed induction generator (DFIG) is the research focus of international experts. The winding fault is up to 12.89 % [1] among the common faults of wind power generator system. After the winding inter-turn short-circuit fault occurs, it causes overheating of the windings and unit vibration intensified and may destroy the DFIG. So it is necessary to detect this fault as early as possible. Research in the area of fault diagnosis of DFIG generates keen interest

L. Zhang (✉) · J. Li · H. Yu
School of Electrical and Electronic Engineering, North China Electric Power University, 071003 Baoding, China
e-mail: dapengniaoroc@163.com

J. Li
e-mail: junqing03@163.com

H. Yu
e-mail: 694878788@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_6

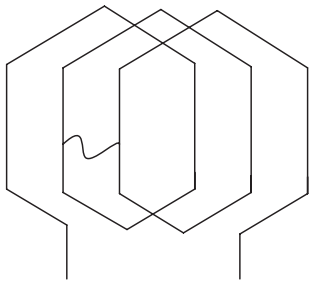


Fig. 6.1 Schematic of winding inter-turn short circuit

as the clamor for renewable energy. Spectral characteristics of voltage and current are used to detect this fault [2], and other methods and detecting data are used to determine whether there is a winding inter-turn short fault [3–6]. But those experts above do not study on how to locate the position of fault slots.

Detection coil has been applied to the detection of synchronous generator rotor turn short-circuit fault [7, 8]. When this fault occurs, the symmetry of the generator is destroyed, so as to the flux density. The effective numbers of turns in fault slots are reduced, so as to the slot leakage flux and the e.m.f. (electromotiveforce) in the detection coil, and with the deepening of the fault, we have the greater of the value of the reduced [9].

With the inter-turn short fault occurs, at least two slots' leakage flux is changed, as shown in Fig. 6.1. In order to locate the fault slot position, detection coil is fixed in the gap of DFIG to collect the change of leakage flux, and the optoelectronic device is coaxially mounted with detection coil to set time reference. In this paper, wavelet decomposition is used to analyze those monitoring data, and then, it is easy to calculate the position of fault slots.

6.2 Fundamental

Wavelet decomposition has advantage over Fourier analysis at time-domain analysis. Moreover, in the field of singular point detection, wavelet decomposition has unmatched advantages over other methods. In this paper, the mutation points and time-domain information should be detected, so wavelet decomposition is the best choice.

Generally, mutant signal in generator fault diagnosis can be divided into two categories: the first discontinuity point and the second type of discontinuity point. The first point of discontinuity, i.e., a certain time, the signal amplitude is mutated, causing signal discontinuities. The second type of discontinuity point, i.e., the amplitude of the signal without mutation, but the occurrence of a mutation appears in order of differential signal and discontinuous [10]. The mutation in this study belongs to the second category discontinuity points.

Let an arbitrary smooth function satisfy the following conditions:

$$\int_{-\infty}^{\infty} \theta(x)\,dx = 1, \quad \lim_{x\to\infty} \theta(x) = 0 \tag{6.1}$$

Select the first derivative of the wavelet function $\psi(x) = \frac{d\theta(x)}{dx}$ and $\psi(x)$ satisfy the conditions. Let $\theta = \frac{1}{s}\theta\left(\frac{x}{s}\right)$, and then, $Wf(x)$ can be represented by $f(x)$.

$$Wf(x) = f(x)\psi_s(x) = s\frac{d}{dx}(f(x)\theta s(x)) \tag{6.2}$$

Therefore, if you choose a smooth function of a first-order differential wavelet, it can detect the mutation points of $f(x)$ by wavelet transform of the maximum amplitude of $Wf(x)$.

N is the vanishing moments of wavelet function dbN, with the larger of the order, the stronger of the ability, which reflects the signal frequency detail of wavelet transform. In the motor fault diagnosis, low-frequency component parts, in which requiring N does not have to be too high, should be concerned more, especially in the analysis of the current signal. In the selection of wavelet basis, its regularity should also be considered. In order to analyze the mutation signals effectively, the wavelet function which with bad regularity is selected. When singularity appears, it contains high-frequency component and it is irregular, so five layers of db3 wavelet decomposition are used in this paper.

6.3 Settings of Simulation

In this paper, simulation model is built in finite element analysis software Ansoft and run the model after setting the necessary conditions. Collect the values of radial flux density when the DFIG is normal, and rotor winding inter-turn short fault occurs.

The simulation model is built based on the real DFIG in the laboratory, and the basic parameters are shown in Table 6.1.

Make the following assumptions in order to simplify the way of solving simulation:

1. Without considering the saturation of ferromagnetic material;
2. Ignoring the outer surface of the stator leakage magnetic field, and the cylindrical of stator flux boundary is parallel;
3. The rotor winding and rotor core insulation;
4. Take radial flux density as analysis objects.

Table 6.1 Basic parameters of the doubly fed induction generator (DFIG)

P_N:	5.5 KW	n_N:	1440 r/min
Outer diameter of the stator	210 mm	Inner diameter of the stator	136 mm
Outer diameter of the rotor	135.2 mm	Stator and rotor slots	36/24
Inner diameter of the rotor	48 mm	Gap	0.4 mm
Shaft length	155 mm	Number of pole pairs	2

6.4 Simulation Results and Analysis

Simulation is set as above, the rotor is running counterclockwise, rotor slots are numbered clockwise, and the inter-turn short fault is set in the 19th slot and 24th slot. The initial time and initial position are specified at the time when No. 1 slot geometric center axle with the x-axis intersects, flux density values are collected clockwise at the same time. The distance between rotor axis and detection coil is radius named r. The circumference of the flux density is $c = 2\pi r$. The average distance between two adjacent slots on the circumference of the sample is $\Delta l = \frac{2\pi r}{Z}$. The rotor speed is named *n*, and cycle time is named $T = \frac{60}{n}$. The time difference when the geometric center of two adjacent slots under detection coil is as follows:

$$\Delta t = \frac{T}{Z} = \frac{60}{n \cdot Z} \tag{6.3}$$

At any time, the 1st slot position is $1_0 = c - c*(\theta/360)$, where θ is the angle of rotation of the rotor. Considering the interaction between each slot flux density, the scope of one slot is extended to each side at the length of $0.5\Delta l$ in the geometric center of the slot.

In accordance with the above provisions, assuming that a fault occurs in the slot whose number is k, where k is a positive integer. *l* represents the position of the maximum or minimum value of flux density difference, so the slot number k satisfies the inequality below:

$$l_0 + (k\text{-}(Z+1))\Delta l - 0.5\Delta l < l < l_0 + (k - (Z+1))\Delta l + 0.5\Delta l \tag{6.4}$$

In this section, the slip $s = 0.2$ is taken as an example, and the initial position and the initial slot numbers of generator are shown in Fig. 6.2.

The simulation model runs at slip of 0.2, and the positions of rotor slots at the time of 0.42 s are shown in Fig. 6.3.

Collect the values of flux density at the entire circumference when the rotor is normal and the rotor winding inter-turn short fault occurs, and make difference of them and then draw the curve of the difference in Fig. 6.4.

In this section, the air gap flux density difference curve at the time of 0.42 s is taken, for example, to make a detail instruction. When the rotor winding inter-turn short fault occurs, the effective turns of the coil in the fault slot are reduced, and the flux density changes significantly at the same time, but others are not. The position of the maximum difference is $l_1 = 151.3\text{mm}$, as shown in Fig. 6.4, and the position of the minimum difference is $l_2 = 244.7\text{mm}$. Then, the effective turns of the coil in the slots which the position mentioned above corresponding to are changed greatly by the analysis above.

The basic data of DFIG at the time of 0.42 s are shown in Table 6.2.

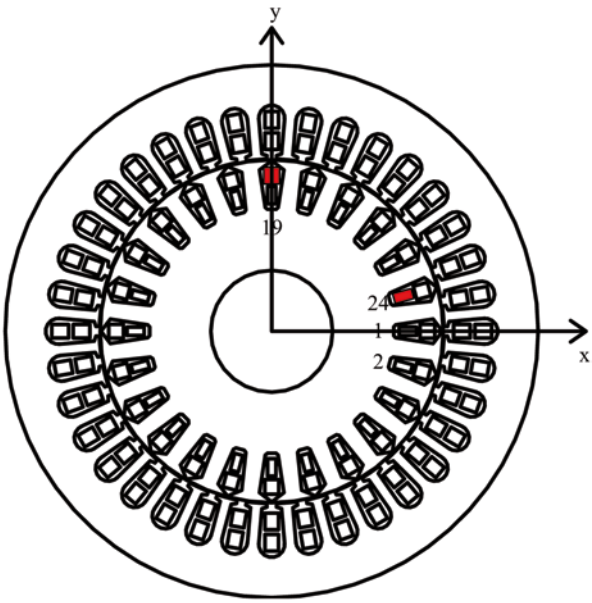


Fig. 6.2 The initial position and slot numbers of simulation model

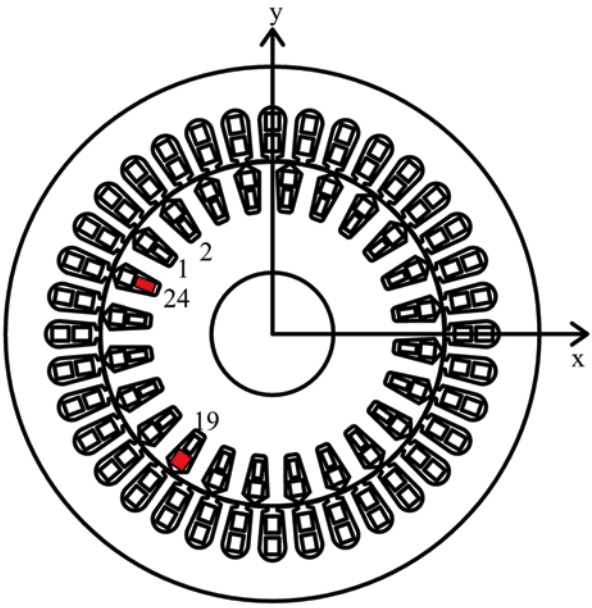


Fig. 6.3 The rotor position at 0.42 s

Taking the basic data into the form to the formula (6.4), the fault slot number k_1 satisfies the inequality below:

$$l_0 + (k_1 - 25)\Delta l - 0.5\Delta l < l_1 < l_0 + (k_1 - 25)\Delta l + 0.5\Delta l \tag{6.5}$$

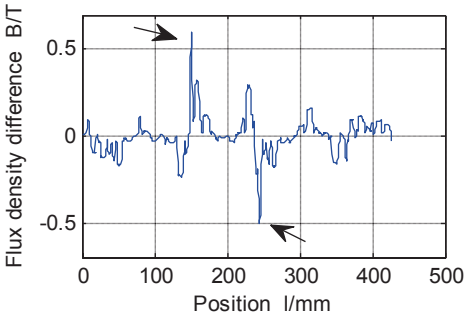


Fig. 6.4 The flux density difference at the time of 0.42 s

Table 6.2 Basic data at the time of 0.42 s

r=67.85mm	c=426.3mm	θ=144.0°
Δl=17.76mm	l_1=151.3mm	l_2=244.7mm

The position of 1st slot: $l_0 = c - c * (144/360) = 255.8\text{mm}$

And the fault slot number k_2 satisfies the inequality below:

$$l_0 + (k_2 - 25)\Delta l - 0.5\Delta l < l_2 < l_0 + (k_2 - 25)\Delta l + 0.5\Delta l \tag{6.6}$$

Then, getting k_1=19 and k_2=24, that is to say the inter-turn short fault occurs in 19th slot and 24th slot which are satisfied with the simulation settings.

6.5 Changes and Analysis of EMF Detection Coil

Whether there is an inter-turn short fault of the DFIG occurs, the degree of the fault can be determined by the spectrum analysis of EMF of detection coil [11].

When the rotor winding is normal, the rate of slot leakage flux changes and the effective turns are equal in each slot. However, the rate of slot leakage flux changes and the effective turns would be reduced when there is an inter-turn short fault, so as to the amplitude of the EMF. Comparing with the normal slots, the EMF amplitude of detection coil conducted by fault slots will be significantly reduced.

To highlight the changes of EMF, the differences of EMF when the rotor winding is normal and the fault is taken as the research subjects.

One complete data are collected, when the generator is running stable, as a sample. Wavelet decomposition is used to detect significant change points, and just d1 is drawn below to show the change points clearly. Combining d1 and the time reference collected by optoelectronic device, we get Fig. 6.5.

In Fig. 6.5, the significant change points are shown clearly, the time when the fault slots are under the detection coil is t_1=37.45ms and t_2=47.85ms. Time error is taken into consideration as the section above, so time error within ±0.5Δt is allowed

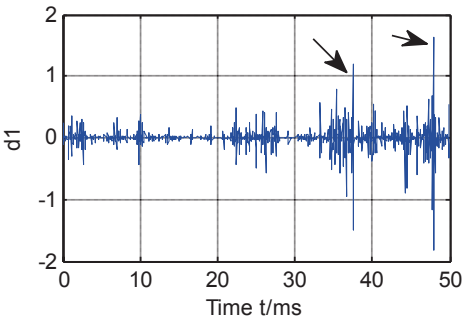


Fig. 6.5 The difference curve of EMF in detection coil

to locate the slots based on different time. So the fault slot number k and fault time t should satisfy the inequality below:

$$(\mathrm{k}-1)\,\Delta t-0.5\Delta t<t<(\mathrm{k}-1)\,\Delta t+0.5\Delta t \tag{6.7}$$

Take (6.3) into (6.7) to calculate the inequality.

The fault slot number k_1 corresponding to the fault time t_1 should satisfy the inequality below:

$$(\mathrm{k}_1-1)\Delta t-0.5\Delta t<t_1<(\mathrm{k}_1-1)\Delta t+0.5\Delta t \tag{6.8}$$

So as to k_2 and t_2 should satisfy the inequality below:

$$(\mathrm{k}_2-1)\Delta t-0.5\Delta t<t_2<(\mathrm{k}_2-1)\Delta t+0.5\Delta t \tag{6.9}$$

After calculating, we get $\mathrm{k}_1=19$ and $\mathrm{k}_2=24$, that is to say the inter-turn short fault occurs in 19th slot and 24th slot which are satisfied with the simulation settings.

A series of different data when $\mathrm{s}=0.1$ and $\mathrm{s}=0.2$ is collected to verify the positioning method proposed in this paper, and the results show that this positioning method is accurate and reliable.

Conclusion

In this paper, the feasibility of locating the fault slots is discussed, and locating the fault slot position by wavelet decomposition has been done successfully. Conclusions of the study are as follows:

1. When the rotor winding inter-turn short fault occurs, the flux density corresponding to the fault slots changes significantly, and the degree of change increases with the degree of the fault.
2. When the fault occurs, the effective turns of coil are reduced, and the leakage flux of the fault slots is reduced, too. Then, the fault slot position is located by

the position to which the maximum and minimum of flux density difference correspond to. Analysis shows that it is feasible to determine the location of the rotor winding inter-turn short fault slots from the perspective of the flux density.
3. When there is the fault, the EMF of detection coil generated by fault slots is decreased. After the calculation of the data, collecting by the combination of detection coil and optoelectronic device, positioning by wavelet decomposition, the fault slots numbers would be determined.

This study provides a positioning method and a theoretical reference for the study of rotor winding inter-turn short fault.

Acknowledgements This work was supported by Natural Science Foundation of Hebei Province in China (E2014502015).

References

1. Wei S, Fu Y, Ma H. Stator winding inter-turn short-circuit diagnosis and experimental research on doubly-fed wind generator. Power Syst Prot Control. 2010;38(11):25–8 (In Chinese).
2. Li J, Wang D, He L. Steady-state analysis on inter-turn short circuit fault of stator windings in doubly-fed induction generators. Autom Electr Power Syst. 2013;37(18):103–7. doi:10.7500/AEPS201301190 (In Chinese).
3. Ma H, Shi W, Han J, Chen J, Chen T. Double-fed induction generator rotor fault diagnosis considering control strategies of rotor-side converters. Proc CSEE. 2013;33(18):119–25 (In Chinese).
4. Qian Y, Ma H. A survey of fault diagnosis method for doubly-fed induction motor. Larg Electr Mach Hydraul Turbine. 2011;5:5–8 (In Chinese).
5. Zhang Z, Ma H, Qian Y, Cheng J. Simulation on inter-turn short circuit faults of stator winding via finite-element method for DFIG. High Volt Appar. 2012;48(8):24–33 (In Chinese).
6. Chen J, Ma H, Shi W, Chen T. Study of stator inter-turn short fault diagnosis of DFIG-based wind turbines based on Hilbert-Huang transform. Larg Electr Mach Hydraul Turbine. 2013;2:34–38 (In Chinese).
7. Sun Y, Yu X, Wei K, Huang Z, Wang X. A new type of search coil for detecting inter-turn faults in synchronous machines. Proc CSEE. 2014;34(6):917–24 (In Chinese).
8. Han L. Dynamic identification for inter-turns short fault in turbogenerator. Baoding, North China Electric Power University. 2011 (In Chinese).
9. Wu Y, Li Y, Li H, Xu B. Preliminary exploration of fault diagnosis for rotor bar breaking of squirrel cage asynchronous motors based on coil detection. J N China Electr Power Univ. 2012;39(3):1–5 (In Chinese).
10. Ren Z, Huang Q, Huang WY. New ways of fault detection in generators based on real wavelet transforms. Proc CSEE. 2000;20(6):58–60 (In Chinese).
11. Shah D, Nandi S, Neti, P. Stator-interturn-fault detection of doubly fed induction generators using rotor-current and search-coil-voltage signature analysis. IEEE Trans Ind Appl. 2009;45(5):1831–42.

Chapter 7
A 3D-Coupled Field-Circuit Model for Analyzing the Internal Short-Circuit Faults of Power Transformer

Guozhao Hu, Xixiu Wu, Fei Long and Shipu Wu

Abstract To characterize the complex physical behavior of power transformer's internal fault is very important for its protection and fault detection. Therefore, this chapter presents a complete three-dimensional-coupled field-circuit model to analyze the internal short-circuit faults of power transformer. The coupled mechanism and modeling which are used to evaluate the performance of transformer under internal short-circuit fault condition are introduced as well. A 500 kVA, three-phase, oil-immersed, core-type transformer is modeled by this field-circuit 3D-coupled method. Based on this, a deep analysis for turn-to-turn and internal phase-to-phase faults is made. The distribution of magnetic field and electrical characteristics which contribute to the development of reliable and sensitive fault detection methods are discussed in this study.

Keywords Power transformer · Internal short circuit · Field-circuit coupled · Model

7.1 Introduction

Power transformer is a major power system component that permits economical power transmission with high efficiency and low series-voltage drops. Statistics shows that nearly 70 % of transformer failures are caused by undetected internal short-circuit faults [1]. Therefore, a thorough understanding of the complex physical behavior of transformer internal faults can provide a sound theoretical basis for

X. Wu (✉) · G. Hu
College of Automation, Wuhan University of Technology, 430070 Wuhan, China
e-mail: wuxixiu@163.com

F. Long
The 712 Research Institute, China Shipbuilding Industry Corporation, 430064 Wuhan, China

S. Wu
Electric Power Research Institute of China, 430074 Wuhan, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_7

developing a reliable and sensitive method to detect the faults. For this reason, it is necessary to establish transformer transient model and analyze the characteristics of transformer under internal short-circuit condition. There are three classical models used to perform the transformer internal fault, which are electric equivalent circuit model [2–3], magnetic circuit model [4], and the 2D finite-element model [5–6]. However, these models are not effective consideration of transformer internal magnetic field and cannot be accurately used to characterize the internal short-circuit faults of power transformer.

In an attempt to obtain more detailed characteristic of transformer under internal faults, a 3D field-electric circuit-coupled model based on time-stepping finite-element method is presented. Several typical internal faults are simulated. Comparing simulating results with the nameplate values, the correctness of this model is validated.

7.2 Field-electric Circuit-Coupled Model

7.2.1 Magnetic Field Equations

Considering the difficulty in meshing and simulation times, the silicon steel is considered as isotropic without hysteresis, and eddy current in the laminated steel is also ignored. Nevertheless, the core loss can be calculated by Bertotti loss separation model [7–8]. Therefore, the partial differential equation of 3D nonlinear magnetic field can be expressed as:

$$\nabla \times \frac{1}{\mu} \nabla \times \boldsymbol{A} = \boldsymbol{J}_s \tag{7.1}$$

where $\boldsymbol{A}$, μ, and $\boldsymbol{J}_s$ are the magnetic vector potential, the magnetic permeability, and the source current density, respectively.

The Bertotti loss separation model can be written as:

$$P_{total} = P_{hys} + P_{cls} + P_{exc} = k_h f B_m^2 + k_c (fB_m)^2 + k_e (fB_m)^{1.5} = k_1 B_m^2 + k_2 B_m^{1.5} \tag{7.2}$$

where P_{total}, P_h, P_{exc}, and P_{cls} are the total loss, hysteresis loss, classical eddy current loss, and excess loss, respectively. f is the frequency of the magnetic induction, and k_h, k_c, and k_e are the hysteresis loss coefficient, the eddy current loss coefficient, and the excess loss coefficient, respectively.

Assuming that the solution region is meshed with tetrahedron elements, Eq. (7.1) can be spatially dispersed based on the edge finite method. The interpolation function is given by

$$\boldsymbol{W}_e = \sum_{\alpha}^{6} \boldsymbol{W}_{ea} A_\alpha = \sum_{\alpha}^{6} (N_i \nabla N_j - N_j \nabla N_i) A_\alpha \tag{7.3}$$

where N_i is the element shape function, A_α is the integral of magnetic vector potential on the element edge, and i, j, and α are the node and edge numbers of the tetrahedron, respectively.

According to Eq. (7.3), the dispersed matrix equation can be obtained as follows:

$$\boldsymbol{KA} = \boldsymbol{F} \tag{7.4}$$

$$\text{With } \boldsymbol{K}^e_{\alpha\beta} = \int_{\Omega_n} \nabla \times \boldsymbol{W}_{e\alpha} \frac{1}{\mu} \nabla \times \boldsymbol{W}_{e\beta}, \; \boldsymbol{F}^e_{\alpha} = \int_{\Omega_n} \boldsymbol{W}_{e\alpha} \cdot \boldsymbol{J}_a d\Omega.$$

where $\boldsymbol{K}$, $\boldsymbol{A}$, and $\boldsymbol{F}$ are the coefficients of stiffness matrix, the vector of magnetic vector potential, and the right side vector, respectively.

According to the transformer geometry, Eq. (7.4) can be decomposed to explain the coupling principle between magnetic field and electric circuit clearly.

$$\boldsymbol{KA} = \sum_{j=0}^{m-1} \boldsymbol{F}_j = \sum_{j=0}^{m-1} \boldsymbol{B}_j I_j \tag{7.5}$$

$$\boldsymbol{B}^e_{\alpha} = \frac{n}{s} \int_{\Omega_n} \boldsymbol{W}_{n\alpha} \cdot l_a d\Omega \tag{7.6}$$

where ***B0*** corresponds to the no current area which is composed of core and tank. ***B1***, ***B2*** … ***Bm-2***, and ***Bm-1*** correspond to the current area which are composed of windings. $\boldsymbol{l}_a, n,$ *and* s are the coil direction vectors which are field unit vector and tangential at the windings, the number of winding turns, and the total cross-sectional area of windings (only including metal region), respectively.

7.2.2 Electric Circuit Equations

The circuit equation of the transformer winding is

$$\boldsymbol{TU} = -\boldsymbol{TE} + \boldsymbol{TRI} + \boldsymbol{TL} \frac{d\boldsymbol{I}}{dt} \tag{7.7}$$

where $\boldsymbol{T}$ is the matrix which represents the connection between each independent winding. U, I, E, L, and R are the terminal voltage sources, the winding current, the back-emf of windings, and the inductance and resistance of each winding, respectively.

7.2.3 Field-Electric Circuit-Coupled Equations

The total induced back-emf in winding can be calculated from the magnetic vector potential:

$$e = -\frac{d\phi}{dt} = -\frac{d}{dt}\oint_c \boldsymbol{A}\cdot d\boldsymbol{l} = -\frac{n}{s}\frac{d}{dt}\int_{\Omega_c} \boldsymbol{A}\cdot \boldsymbol{l}_\alpha d\Omega \tag{7.8}$$

Then, the Eq. (7.7) becomes

$$\boldsymbol{TU} = \boldsymbol{TB}^{*T}\frac{d\boldsymbol{A}}{dt} + \boldsymbol{TRI} + \boldsymbol{TL}\frac{d\boldsymbol{I}}{dt} \tag{7.9}$$

Where $\boldsymbol{B}^{*T} = \begin{bmatrix}\boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_{m-2} & \boldsymbol{B}_{m-1}\end{bmatrix}^T$.

From (7.4) and (7.9), the transient field-electric circuit-coupled equations can be obtained as

$$\boldsymbol{GX} + \boldsymbol{P}\dot{\boldsymbol{X}} = \boldsymbol{Q} \tag{7.10}$$

where $\boldsymbol{G} = \begin{bmatrix}\boldsymbol{K} & -\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{TR}\end{bmatrix}; \boldsymbol{X} = \begin{bmatrix}\boldsymbol{A} \\ \boldsymbol{I}\end{bmatrix}; \boldsymbol{P} = \begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{TB}^{*T} & \boldsymbol{TL}\end{bmatrix}; \boldsymbol{Q} = \begin{bmatrix}0 \\ \boldsymbol{TU}\end{bmatrix}$.The Eq. (7.10) will be dispersed by the Crank–Nicholson method in the time domain, and the dispersed equations are nonlinear which can be solved by Newton–Raphson method.

7.3 Analysis and Discussion

7.3.1 Transformer Parameters and Modeling

Based on the mathematical principles discussed above, a three-phase, oil-immersed, core-type transformer has been simulated under several typical internal fault cases. Transformer parameters are shown in Table 7.1.

Figure 7.1 presents the simulation schematic diagram of this transformer. As shown in this figure, the principle used for modeling internal winding fault in field

Table 7.1 Transformer parameters

Rated power	500 kVA	Connection	Yyn0
HV-rated voltage/current	6.6 kV/43.8 A	Short-circuit impedance	3.95 %
LV-rated voltage/current	400 V/721.7 A	Winding type	Layer
Rated frequency	50 Hz	Core loss	727 W

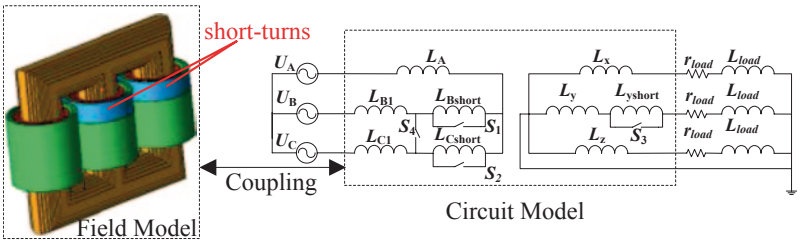


Fig. 7.1 Field-electric circuit-coupled schematic diagram for simulation

domain is to divide the winding in which the internal fault occurs into two subcoils: the short-circuit part and the remaining coils. Corresponding to Eq. (7.7), the electric circuit equations of each independent winding can be written as

$$\begin{cases} u_A = -e_A + R_A i_A \\ u_{B_1} = -e_{B_1} + R_{B_1} i_{B_1} \\ u_{C_1} = -e_{C_1} + R_{C_1} i_{C_1} \\ u_x = -e_x + R_x i_x + r_{load} \boldsymbol{i}_x + L_{load} \dfrac{di_x}{dt} \\ u_y = -e_y + R_y i_y + r_{load} i_y + L_{load} \dfrac{di_y}{dt} \\ u_z = -e_z + R_z i_z + r_{load} i_z + L_{load} \dfrac{di_z}{dt} \\ u_{Bshort} = -e_{Bshort} + R_{Bshort} i_{Bshort} \\ u_{Cshort} = -e_{Cshort} + R_{Cshort} i_{Cshort} \end{cases} \tag{7.11}$$

In the circuit, the switch is closed for transformer internal fault simulation which can be represented by the alteration of transformation matrix. For example, to model the turn-to-turn short-circuit fault of B-phase in this transformer, the transformation matrix can be written as follows:

Normal operation condition turn to turn fault of B - phase

$$T_1 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 1 & -1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad T_2 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{7.12}$$

To multiply Eq. (7.11) by Eq. (7.12), the electric circuit equations of the transformer under different operation conditions are as follows.

Normal operation condition:

$$\begin{cases} u_A - u_B = u_A - (u_{B_1} + u_{Bshort}) = -e_A + R_A i_A + (e_{B_1} + e_{Bshort}) - (R_{B_1} i_{B_1} + R_{Bshort} i_{Bshort}) \\ u_B - u_C = (u_{B_1} + u_{Bshort}) - u_C \\ \qquad = -(e_{B_1} + e_{Bshort}) + (R_{B_1} i_{B_1} + R_{Bshort} i_{Bshort}) + (e_{C_1} + e_{Cshort}) - (R_{C_1} i_{C_1} + R_{Cshort} i_{Cshort}) \\ u_C - u_A = (u_{C_1} + u_{Cshort}) - u_A = -(e_{C_1} + e_{Cshort}) + (R_{C_1} i_{C_1} + R_{Cshort} i_{Cshort}) + e_A - R_A i_A \\ u_x = -e_x + R_x i_x + r_{load} i_x + L_{load} \dfrac{di_x}{dt} \\ u_y = -e_y + R_y i_y + r_{load} i_y + L_{load} \dfrac{di_y}{dt} \\ u_z = -e_z + R_z i_z + r_{load} i_z + L_{load} \dfrac{di_z}{dt} \end{cases} \tag{7.13}$$

Turn-to-turn fault of B-phase:

$$\begin{cases} u_A - u_B = u_A - (u_{B_1} + u_{Bshort}) = -e_A + R_A i_A + (e_{B_1} + e_{Bshort}) - (R_{B_1} i_{B_1} + R_{Bshort} i_{Bshort}) \\ u_B - u_C = u_{B_1} - u_C = -e_{B_1} + R_{B_1} i_{B_1} + (e_{C_1} + e_{Cshort}) - (R_{C_1} i_{C_1} + R_{Cshort} i_{Cshort}) \\ u_C - u_A = (u_{C_1} + u_{Cshort}) - u_A = -(e_{C_1} + e_{Cshort}) + (R_{C_1} i_{C_1} + R_{Cshort} i_{Cshort}) + e_A - R_A i_A \\ u_x = -e_x + R_x i_x + r_{load} i_x + L_{load} \dfrac{di_x}{dt} \\ u_y = -e_y + R_y i_y + r_{load} i_y + L_{load} \dfrac{di_y}{dt} \\ u_z = -e_z + R_z i_z + r_{load} i_z + L_{load} \dfrac{di_z}{dt} \\ u_{Bshort} = -e_{Bshort} + R_{Bshort} i_{Bshort} = 0 \end{cases} \tag{7.14}$$

7.3.2 Turn-to-Turn Short-circuit Fault

First, a turn-to-turn fault involving 5.6% of the turns (20 turns) that occurs at $t = 0.04$ s on the primary windings of phase "B" is simulated. The primary windings are supplied by the rated voltage, and the secondary windings are connected to rate resistive inductive load (power factor is 0.95).

Figure 7.2 presents the terminal currents and circulating current in the shorted turns. When a turn-to-turn fault occurs on the primary winding, a very large circulating current will flow in the shorted turns. The circulating current is the source of an MMF, which opposes to the primary winding and shares the same direction with the secondary. Thus, the inter-turn fault has an effect in decreasing the fundamental flux, thus consequently increasing the primary current and decreasing the secondary current to compensate it.

Table 7.2 Simulation results compared with nameplate values

Parameters	Nameplate values	Simulation results	Error
HV voltage (kV)	6.6	6.6	0
LV voltage (V)	400	391.9	−2.03 %
HV current (A)	45.82	43.8	−4.41 %
LV current (A)	721.7	707.11	−2.02 %
Power (KVA)	500	491	−1.80 %
Core loss(W)	727	692	−4.81 %
Short-circuit impedance[a]	3.95 %	3.81 %	−3.42 %

[a] The short-circuit impedance was calculated using energy method [9]

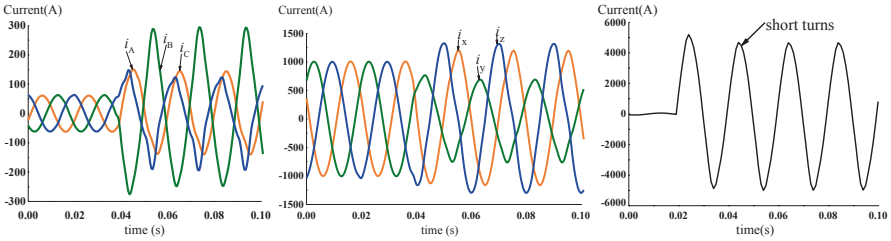


Fig. 7.2 Current waveforms for the turn-to-turn fault on the primary winding of transformer, occurring at the $t = 0.04$ s of phase B.

Table 7.2 shows the comparing results between calculating values under normal operation condition and the nameplate values. From this table, we know that the calculation errors are within 5 %, which indicates that the transient operating state of transformer can be simulated accurately by this model.

Figure 7.3 presents the flux density distribution inside the transformer under normal operation conditions. It clearly reveals that the flux density distribution in the healthy transformer has a horizontal symmetry axis that passes through the center of the transformer core limbs and the magnetic flux in the core is much higher than the leakage flux.

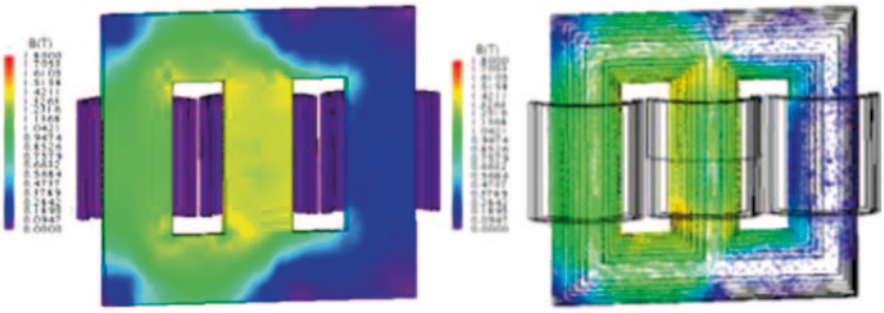


Fig. 7.3 Flux density distribution inside the transformer under normal operating conditions

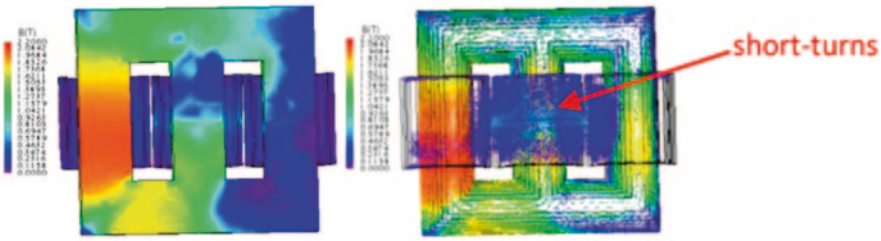


Fig. 7.4 Flux density distribution inside the transformer under turn-to-turn fault conditions

The distribution of magnetic flux density after the occurrence of turn-to-turn fault is shown in Fig. 7.4.

Compared with the normal operation condition, the core gets into saturation, the leakage flux around shorted turn increases significantly, and the flux in the core limb surrounded by the shorted turn decreases; the symmetry in the magnetic flux density distribution does not exist.

When a turn-to-turn fault occurs on the secondary winding, the current and flux density distribution has similar characteristics with the occurrence on the primary. Such as the primary current and the secondary current of non-fault phase will increase, the secondary current of fault phase will decrease, and the magnetic flux density distribution is asymmetrical inside the faulty transformer. Due to the limited space, the current waveform and flux density distribution are not shown here.

7.3.3 Internal Phase-to-Phase Short-circuit Fault

Since the inside windings that are closed to core are the secondary, the primary windings are outside. Winding internal phase-to-phase short-circuit fault that happens on the same longitudinal position of the B, C primary windings is analyzed in this study. Each phase involves 12 shorted turns. Terminal current and circulating current in shorted turns are shown in Fig. 7.5. Figure 7.6 presents the flux density distribution after the fault occurrence.

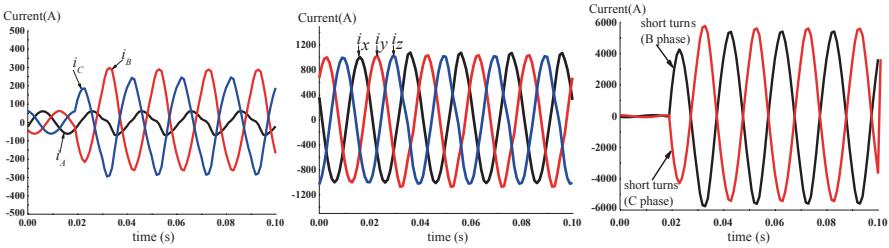


Fig. 7.5 Current waveforms for internal phase-to-phase fault on the primary winding of transformer, occurring at the $t = 0.04$ s of phase *B*, *C*

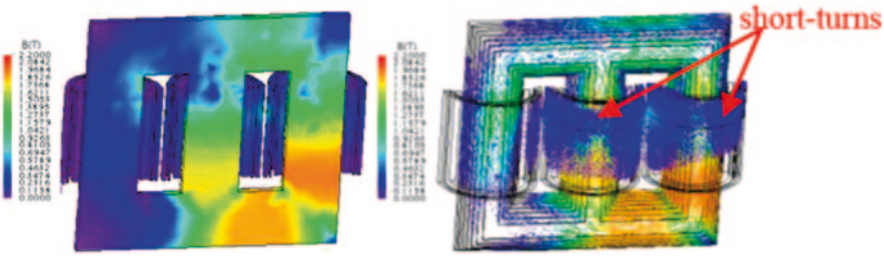


Fig. 7.6 Flux density distribution inside the transformer under internal phase-to-phase fault conditions

In the figure, the current on the primary side of phase B, C increases evidently, but the current of phase A as well as secondary current remain almost unchanged. Currents flowing in the shorted turns are very large and phase opponent in phase B, C. The flux density distribution characteristic shows familiar to the turn-to-turn fault discussed above.

7.4 Conclusion

In this paper, a 3D-coupled field-circuit model has been developed and applied in analyzing the internal short-circuit faults of power transformer. The detailed formulation and modeling to evaluate the performance of transformer under the internal short-circuit fault have been introduced. Field-electric circuit-coupled model provides a direct visual observation of electromagnetic field distribution and change in the power transformer. It is useful for understanding the transient feature of the transformer under different operating conditions. Turn-to-turn fault and internal phase-to-phase fault have been discussed in this study. Some useful characteristics of internal fault have been extracted. These results will help to develop reliable and sensitive methods for detection of the internal faults on the windings for the power transformers.

Acknowledgements This work was supported by National Natural Science Foundation of China (51107093), State Key Laboratory of Advanced Electromagnetic Engineering and Technology (Huazhong University of Science and Technology), and Fundamental Research Funds for the Central Universities (2014-IV-143).

References

1. Abed NY, Mohammed OA. Modeling and characterization of transformers internal faults using finite element and discrete wavelet transforms. Magn IEEE Trans. 2007;43(4):1425–8.
2. Patrick B, Pierre B, Meunier M. A transformer model for winding fault studies. Power Delivery, IEEE Trans. 1994;9(2):690–699.

3. Mladen K, Guo Y. Modeling and simulation of the power transformer faults and related protective relay behavior. Power Delivery, IEEE Trans. 2000;15(1):44–50.
4. Hao WB, et al. A new research for internal fault simulation model of transformer. Power System Technology, 2006. PowerCon 2006. International Conference on. IEEE. 2006;1–6.
5. Wang H, Butler KL. Finite element analysis of internal winding faults in distribution transformers. Power Delivery IEEE Trans. 2001;16(3):422–8.
6. Behjat V, Vahedi A. Numerical modelling of transformers interturn faults and characterising the faulty transformer behaviour under various faults and operating conditions. IET Electr Power Appl. 2011;5(5):415–431.
7. Bertotti G. Hysteresis in magnetism: for physicists, materials scientists, and engineers[M]. New York: Academic Press Inc; 1998. pp. 10–548.
8. Sergey EZ, et al. Static and dynamic hysteresis models for studying transformer transients. IEEE Trans on Power Delivery. 2011;26(4):2352–62.
9. Zhang J, et al. Hurley Leakage inductance calculation for planar transformer with magnetic shunt. Energy Conversion Congress and Exposition (ECCE), IEEE. 2013. pp. 643–648.

Chapter 8
Position Detection and Error Compensation for Tubular PMLSM with Hall-effect Sensors

Yong Zhang, Fang Li and Peiqing Ye

Abstract In this chapter, a new position detection method for tubular permanent magnet linear synchronous motor (TPMLSM) is proposed. Instead of using grating sensors, the linear hall-effect sensors are adopted, which reduce the cost largely. In some special PMLSM, the position of hall-effect sensor is restrained by the size of motors, leading to an extra error. So a predetermined lookup table for error compensation is proposed in this paper. Furthermore, the finite element analysis based on Ansoft also demonstrates the validity and reliability of the position detection and error compensation of the tubular PMLSM based on hall-effect sensors.

Keywords Tubular PMLSM · Hall-effect sensors · Position detection · Error compensation · Lookup table

8.1 Introduction

In various industrial applications, the direct linear servo systems with tubular PMLSM are becoming more and more widespread [1]. And compared to traditional systems, that kind of linear servo systems have many advantages of faster transient response, control capabilities, and simpler manufacture due to tubular PMLSM, making it a more attractive transmission mode [2]. In traditional systems, the expensive grating sensors, which cost more than some hundred dollars, are generally utilized in position control that increases the whole cost largely [3]. But in some application fields, the requirement of position accuracy is not high enough to use the expensive grating sensors [4], so a new method of position detecting based on a kind of low-cost hall-effect sensors has been proposed and studied in this paper. Compared to grating sensors, hall-effect sensors only cost less than five dollars,

Y. Zhang (✉) · F. Li · P. Ye
The State Key Laboratory of Tribology, Tsinghua University, 100084 Beijing, China
e-mail: zhangy10@126.com

F. Li
e-mail: lifphd@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_8

so the cost of the PMLSM is reduced largely. And by using hall-effect sensors, the measurement process of the position detecting is simplified greatly. However, to pursue a higher accuracy, the predetermined lookup table is made to compensate the position errors.

The remainder of this chapter is organized as follows. First, the topology of tubular PMLSM and the theoretical analysis for position detection are illustrated in Sect. 2. In Sect. 3, the model of PMLSM is built and the finite element analysis based on Ansoft is finished. Then, the experiment results are illustrated in Sect. 4. In Sect. 5, error compensation for PMLSM is finished based on finite element analysis and experiment results. Finally, Sect. 6 concludes the entire paper.

8.2 The Position Detection of Tubular PMLSM with Hall-Effect Sensors

Figure 8.1 shows the topology of tubular PMLSM with two linear hall-effect sensors which are installed in the middle of the stator. As known, the magnet field at the position between mover and stator mainly depends on the distance to the mover. And in an optional position, the magnetic flux density along x-axis can be described as a sinusoidal wave. The linear hall-effect sensors are used to measure the magnetic flux density based on Hall Eeffect. In other words, the output voltage of the sensors is linear to the magnetic flux density. And to detect the position of the mover accurately, the sensors must be apart 90° of electrical angle, which means half polar pitch. In this case, the output signal of the sensors will become sinusoidal wave and cosinusoidal wave. Through anti-trigonometric, the electrical angle can be calculated easily, and then, the position of the mover can be calculated [5].

The output voltage of two hall-effect sensors is U1 and U2, and then,

$$\begin{aligned} U_1 &= A\sin\theta \\ U_2 &= A\sin(\theta + 90^\circ) = A\cos\theta. \end{aligned} \tag{8.1}$$

From the formulas, the electrical angle of the mover θ can be calculated.

$$\theta = \arctan(U_1 / U_2). \tag{8.2}$$

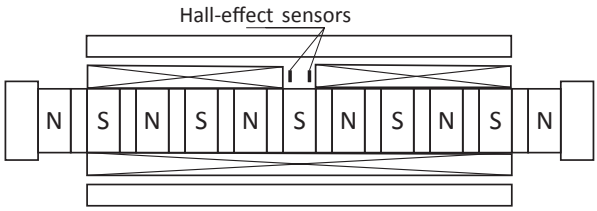


Fig. 8.1 Topology of tubular PMLSM

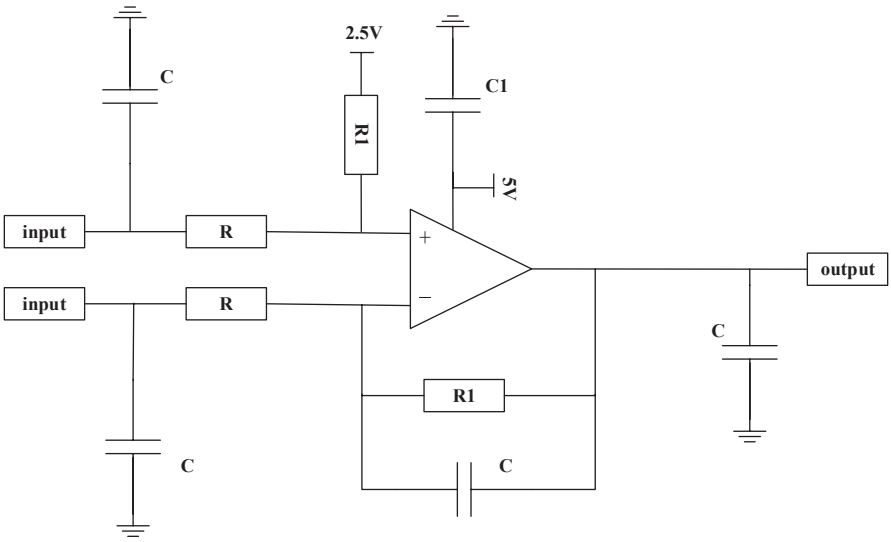


Fig. 8.2 The low-pass filtering and amplifying circuit

And because 90° of electrical angle means half polar pitch (λ), the electrical angle is linear to the distance compared to the initial position of the mover.

$$\theta = \frac{90^{\circ}}{0.5 \times \lambda} \times x = \frac{180^{\circ}}{\lambda} \times x. \tag{8.3}$$

However, in some application fields, the distance between mover and stator is usually decreased to achieve a better performance of thrust, which limits the installation of the hall-effect sensor, making it more difficult to detect the position and decreasing the accuracy. When the accuracy does not meet the requirement, the error compensation is needed. And the position detection and error compensation of the special tubular PMLSM is studied in Sect. 3 and Sect. 5.

Actually, the direct output from sensors is not as good as expecting. Owing to the external factors, such as surrounding temperature and machining errors, so there will be some ripples in the output, which reduces the stability of position control. To restrain the ripple, a low-pass filtering and amplifying circuit as shown in Fig. 8.2 is designed.

8.3 Finite Element Analysis of PMLSM

The model of the tubular PMLSM is built in Ansoft which is a software based on Maxwell differential equations. Then, the finite element analysis is carried out on this model. The magnetic flux density is shown in Fig. 8.3. What we concern the

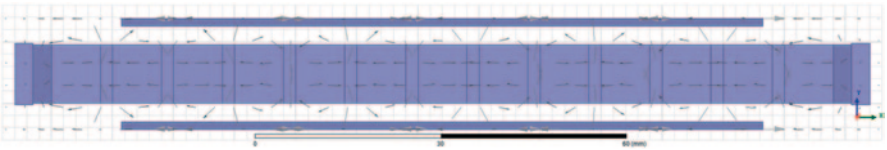


Fig. 8.3 The distribution of magnetic flux density

most is the magnetic flux density on the mounted position of the two sensors, so the input of the sensors is shown in Fig. 8.4. From Fig. 8.4, we can conclude that the signal is sinusoidal wave approximately [6]. But as said it was mentioned before, the decreasing of the distance between mover and stator, which restrains the installation of the hall-effect sensor, leads to a large error compared to the optional mounted position. To visualize the error, the results are shown in Fig. 8.5, where the sine and cosine signals are plotted in both axis [7]. Theoretically, Fig. 8.5 should be a perfect circle, instead of a rhombus.

From Fig. 8.5, we can get the electrical angle, and then, the mover's position can be achieved according to equation (3). From simulation, we can find that the electrical angle of magnetic flux density is approximately linear to the position of the mover, correlating with the theory. The results are shown in equation (4).

$$\theta = 18.0345 \times x - 0.7857 \tag{8.4}$$

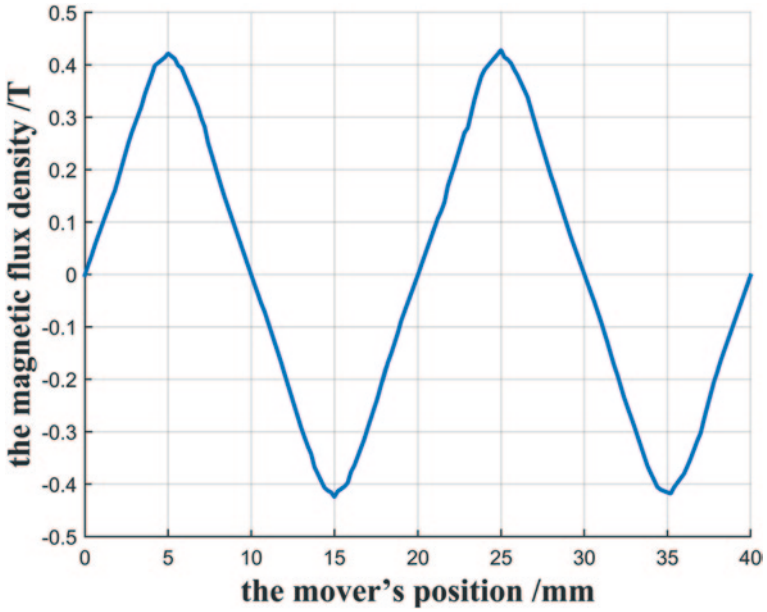


Fig. 8.4 The input magnetic flux density of sensors along the mover's position

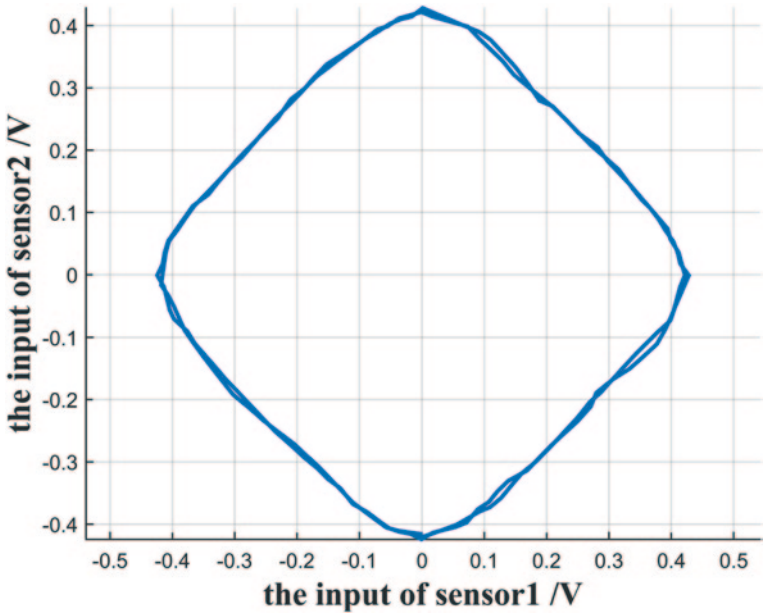


Fig. 8.5 The x–y diagram of the input magnetic flux density

8.4 Experiment Results

Using the linear grating sensors, we carry out an experiment by moving the mover at a fixed speed. Figure 8.6 shows the experiment table and our tubular PMLSM. The output of the low-pass filtering and amplifying circuit is also collected to calculate the position of the mover. To compare with the result of finite element analysis, the output voltage is also shown in Fig. 8.7, and the x–y diagram is shown in Fig. 8.8. We can find that it is in accordance with the finite element analysis. From the data, the electrical angle along the mover's position can be calculated, and the result is as follows:

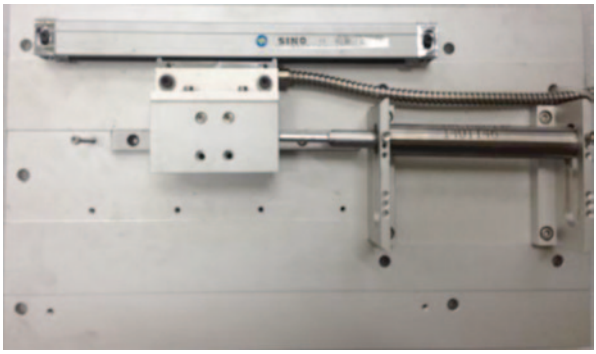


Fig. 8.6 Experimental picture

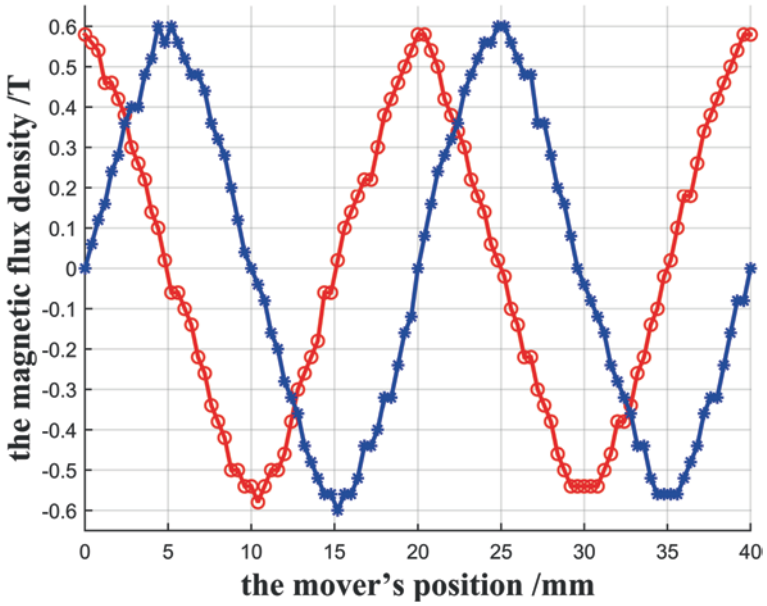


Fig. 8.7 The output of the amplifying circuit along the mover's position

$$\theta = 18.0987 \times x - 2.6724. \tag{8.5}$$

8.5 The Error Compensation of PMLSM Based on Finite Element Analysis and Experiment

From the analysis in Sect. 2, the electrical angle is linear to the distance of the mover. However, the hall-effect sensors do not install on the optional position because of the limitation of the small gap, so there is a theoretical error shown in Fig. 8.9, and from Fig. 8.9, the simulation is up to about 340 μm. Similarly, the actual error is about 500 μm in Fig. 8.9, and two curves have similar changing trend, meaning that the results of simulation agree well with experiment data. However, the theoretical error is much smaller and more regular than the actual error. This is because that the actual parameters of the permanent magnet and linear hall-effect sensors are incompletely consistent with the theoretic one. These differences are mainly caused by the production of PMLSM. Firstly, the hall-effect sensors could be incompletely perpendicular to y-axis, and the distance between two sensors must exist a deviation from half polar pitch, which all contribute to the error.

In order to reduce the error, a lookup table for error compensation which is used in the drivers of PMLSM is made. And after compensation, the error is reduced to 100 μm, and the curve is also shown in Fig. 8.9. In traditional method for position

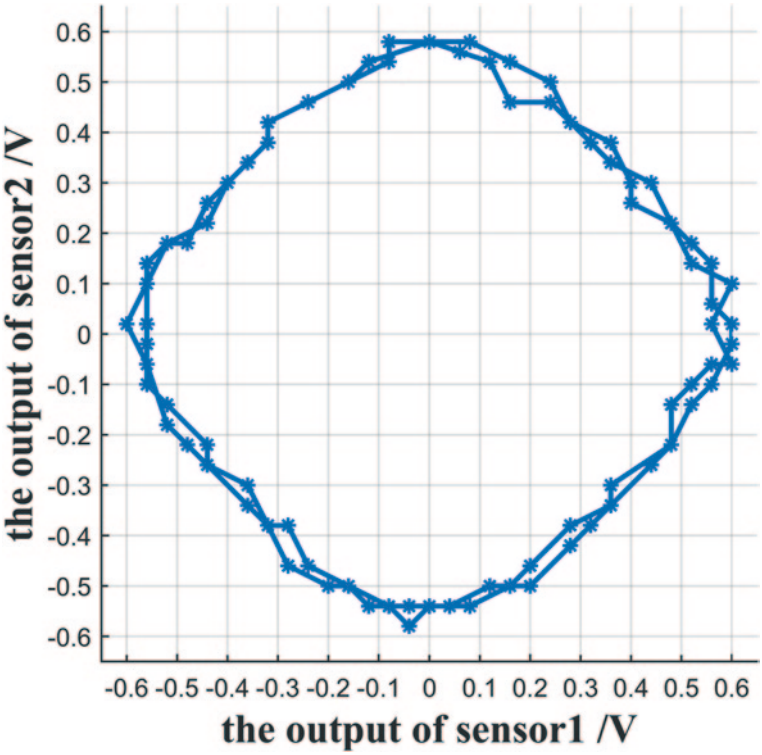


Fig. 8.8 The x–y diagram of the measured signals

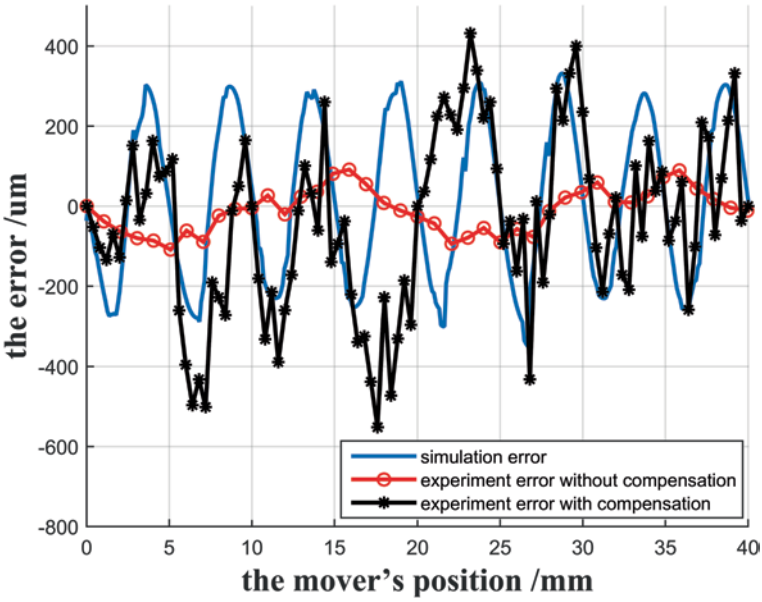


Fig. 8.9 Error curve of the detection angle

detection using grating sensors, the error is about 10 μm. So compared to hall-effect sensors, the grating sensors have higher accuracy, but in some application fields, the accuracy does not need to reach to 10 μm, and using hall-effect sensors, it is enough. So in that kind of application field, it is a better choice to use hall-effect sensors considering the cost.

8.6 Conclusion

In this paper, the finite element analysis and the experiment results all prove that it is possible to use low-cost and available hall-effect sensors to detect the mover's position. And the experiment results we achieved are consistent with the theoretical analysis and simulation results, which are shown from the error curves. The best advantage of the method is the low cost of hall-effect sensors compared to using expensive grating sensors. In the production of PMLSM, there are a lot of machining errors compared to simulation, and in some special PMLSM, the mounted position of hall-effect sensors also decreases the position accuracy. So in order to improve the accuracy in comparison with grating sensors, a lookup table is made to decrease the error from 500 μm to 100 μm and improve the reliability in positioning PMLSM.

Acknowledgements This study was supported by Tribology Science Fund of State Key Laboratory of Tribology of China under Grant No.SKLT12A03, Science and Technology Planning Program of Beijing, China under Grant No. Z141100000514015, and Tsinghua University Initiative Scientific Research Program of China under Grant No.20111081026.

References

1. Liu X, Ye Y, Zheng Z. Study of the linear hall-effect sensors mounting position for PMLSM. 2007s IEEE Conference on industrial electronics and applications; IEEE; 2007; pp. 1175–1178.
2. Rui K, He P, Kenji S, et al. Study of intermittent stator PMLSM control method by a position detector using hall element. 2013 International Conference on electrical machines and systems (ICEMS); IEEE; 2013; pp. 1981–1984.
3. Zheng Xu Z, Tiecai Li T, Yongping Lu Y, et al. Position-measuring error analysis and solution of hall sensor in pseudo-senseless PMSM driving system. The 29th Annual Conference of the IEEE; IEEE; 2003; pp. 1337–1342.
4. Ralf W, Florian S, Junge C, et al. Low cost position sensor for permanent magnet linear drive. IEEE; pp. 1367–1371; 2007.
5. Liu X, Zhuo Z, Ye Y, et al. Position detecting for the air-cored TPMLSM with linear Hall-effect sensors. International Conference on electrical machines and systems, IEEE; 2008; pp. 1417–1420.
6. Liu X, et al. Magnetic field and performance analysis of a tubular permanent magnet linear synchronous motor applied in elevator door system. J Zhejiang Univ SCI A. 2008;9(4): 572–576.
7. Kim S-Y, Lee K, Lee W. An improved rotor position estimation with vector-tracking observer in PMSM drives with low-resolution hall-effect sensors. IEEE Trans Ind Electron. 2011;58(9):4078–4086.

Chapter 9
Application of Ensemble Empirical Mode Decomposition on Stator Inter-turn Short-Circuit Fault in Doubly Fed Induction Generators

Junqing Li, Haibo Yu and Lipeng Zhang

Abstract The stator winding inter-turn short-circuit fault is one of the common faults of doubly fed induction generator (DFIG). In this paper, the multi-loop theory is used for the mathematical modeling, while different faults under different conditions are simulated by MATLAB. Due to the drawbacks in the EMD decomposition method, the results of decomposition cannot distinguish the inter-turn short-circuit fault from some non-fault conditions, while, by introducing the EEMD decomposition method, the problem can be solved. Using the new method, the voltage imbalance, variable rotor speed, and other factors on the interference in the process of inter-turn short-circuit fault diagnosis can be ruled out; therefore, the result of the current spectrum analysis is more reliable. EEMD method can provide a more reliable guarantee for stator inter-turn short-circuit fault detection.

Keywords Doubly fed induction generator (DFIG) · Multi-loop · EEMD · Inter-turn

9.1 Introduction

Currently, the doubly fed induction generator (DFIG) is one of the main equipments of the operational wind power generators; however, its work environment is so bad that it has a high failure rate. Stator inter-turn short-circuit fault accounts for more than 30 % of the generator faults [1]. Therefore, the monitoring and early diagnosis

H. Yu (✉) · J. Li · L. Zhang
School of Electrical and Electronic Engineering, North China Electric Power University, 071003 Baoding, China
e-mail: 694878788@qq.com

J. Li
e-mail: junqing03@163.com

L. Zhang
e-mail: dapengniaoroc@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_9

for DFIG fault play an important role in maintaining the safe and stable operation of the generator and power grid.

9.2 EMD Decomposition and EEMD Decomposition

Empirical mode decomposition (EMD) is part of Hilbert–Huang transform (HHT) which is put forward by Norden E. Huang, which includes two steps: First is decomposing all the target signals into intrinsic mode function (IMF), and the decomposition process is called EMD, then transforming the IMF by Hilbert.

9.2.1 EMD Decomposition

Only the data that meet the following two conditions can be decomposed by EMD: 1 The sum of envelopes of its maximum points and minimum points is symmetric along the timeline; and 2 the number of the extreme value point and the number of passing zero should be equal or differ at most of 1.

One of the disadvantages of EMD method is as follows: Some scale will be lost when the data are not pure white noise; therefore, modal aliasing phenomenon came into existence. It means that an IMF component includes a signal of big scale differences, or a similar scale signal appears in a different IMF component. All data contain the signal and noise in reality life; therefore, the EMD modal aliasing phenomenon is inevitable. This is one of the main drawbacks of EMD [2].

9.2.2 EEMD Decomposition

Ensemble empirical mode decomposition (EEMD) is a data analysis method for noise auxiliary aiming at overcoming the drawback of EMD method [3].

Decomposition steps of EEMD decomposition method are as follows:

1. A new signal $x(t)$ is obtained through putting a group of white noise $\omega(t)$ into original signal;
2. IMF component r_n is obtained by EMD decomposition of new signal;
3. Put different white noise into the target signal and repeat the above two steps;
4. The IMF component groups $c_n(t)$ are obtained after decomposing;
5. Take the corresponding average IMF as the final IMF group;
6. Take the corresponding average of residual component as the final IMF group.

Then, the final signal $x(t)$ equation is as follows: $x(t) = \sum_{n=1}^{m} c_n(t) + r_m(t)$.

Based on the above steps and formulas, EMD and EEMD methods are simulated by MATLAB programming.

9.3 The Mathematical Model of Multi-Loop

The multi-loop mathematical model under normal circumstances can be written as follows [4]:

$$\mathbf{U} = \mathbf{p}\psi + \mathbf{RI} \tag{9.1}$$

where $\boldsymbol{U}$ represents the stator and rotor voltages, $\boldsymbol{I}$ represents the current column vector, respectively, $\boldsymbol{R}$ means the stator and rotor resistances, and p stands for the differential operator.

Flux equation is given as (9.2).

$$\psi = \mathbf{LI} \tag{9.2}$$

$\boldsymbol{L}$ is a matrix made up of self-inductance and mutual inductance of stator winding and rotor winding.

According to the loop, the transformation matrix $\boldsymbol{H}$ of branch and loop under normal circumstances can be written as follows:

$$\mathbf{H} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & \mathbf{0} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \tag{9.3}$$

$\boldsymbol{H}$ and ψ will be substituted into (9.1). A new equation is shown as follows:

$$\mathbf{HU} = \mathbf{H}_p\mathbf{LI} + \mathbf{HL}_p\mathbf{I} + \mathbf{HRI} \tag{9.4}$$

The current in branch and loop of stator and rotor should be as follows:

$$\boldsymbol{I} = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{I}' \tag{9.5}$$

$\boldsymbol{I}'$ is the current of the stator and rotor.

After combining (9.5) with (9.4) and simplifying, the following Eq. (9.6) can be obtained:

$$p\boldsymbol{I}' = (-\boldsymbol{L}')^{-1}\boldsymbol{R}'\boldsymbol{I}' + (\boldsymbol{L}')^{-1}\boldsymbol{U}' \tag{9.6}$$

where $\boldsymbol{U}^{'} = \boldsymbol{HU}, \boldsymbol{L}^{'} = \boldsymbol{HLH}^{\mathrm{T}}, \boldsymbol{R}^{'} = p\boldsymbol{L}^{'} + \boldsymbol{HRH}^{\mathrm{T}}$.

After the transformation, a generator multi-loop mathematical model has been created. Faults are simulated by MATLAB programming, and Eq. (9.6) is solved through using fourth-order Runge-Kutta method, and then each loop current of stator and rotor can be obtained. After transforming using Eq. (9.5), each branch current can be obtained. Short-circuit turns form a new loop when the stator inter-turn short-circuit fault occurs. Thus, considering the new loop in the analysis and calculation, various physical quantities of different degrees of failure can be obtained through the same way.

9.4 Analysis Simulation Results

Based on the above theory, a 5.5-kW doubly fed induction generator is simulated. The generator's basic parameters are as follows: Rated frequency is 50 Hz, rated voltage is 380 V, the stator slot number is 36, the rotor slot number is 24, the pole number is 2, number of turns of each stator coil is 37, number of turns of each rotor coil is 12, and the rotor speed is 1200 rpm.

Line current I_b is constituted by phase current B and C; I_c is constituted by phase current A and C; I_a is constituted by phase current A and B.

Stator and rotor three-phase line current waveforms is shown in Figs. 9.1, 9.2.

As shown in figures, when the motor runs normally, three-phase current of stator is equal, phase difference is 120°, and waveform of rotor is smooth; when 5-turn

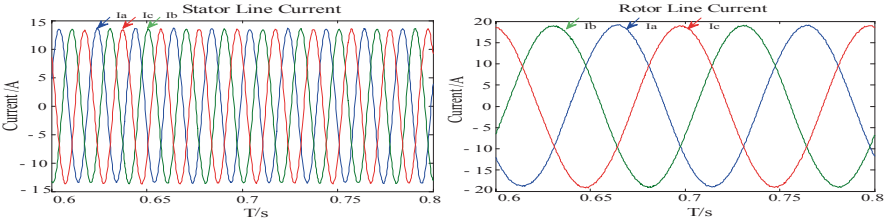


Fig. 9.1 The normal current

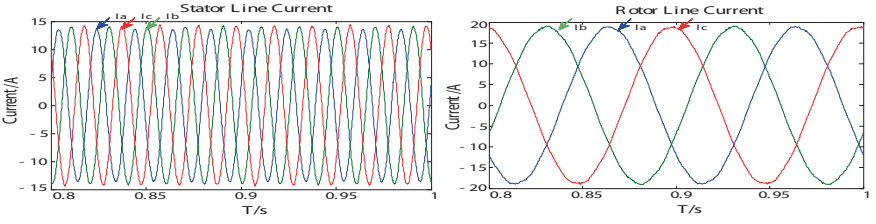


Fig. 9.2 5-turn short circuit of stator C-phase

short-circuit fault occurs, three-phase current of stator will be asymmetrical, and amplitude of current I_b and I_c increases at varying degrees. Because of the influence of harmonic, there are different degrees of volatility in the waveform of rotor three-phase current, but the three-phase current basically keeps symmetry. The simulation results are based on the generator's design principle.

9.4.1 Comparison of Results of EMD and EEMD Analysis

The multi-loop mathematical model simulation results are processed by the EMD and EEMD. Generator reaches steady state at 0.6 s, and the 2-turn short circuit occurs at 0.8 s. Analysis results of stator current I_c are shown in Figs. 9.3, 9.4 below:

The waveforms that start at 0.6 s are EMD and EEMD transformation of short-circuit fault current I_c, and the fault occurs at 0.8 s. IMF 1-2 in Fig. 9.3 shows obvious changes at 0.8 s, and a whole signal waveform is divided into two parts by the EMD method. Due to amplitude, phase and frequency change after the fault occurs; thus, fault waveforms are identified and divided into two segments by EMD decomposition. As known from simulation that 0.8 s is the point of fault, EMD method can identify the time point and is helpful to further determine the generator fault. Figure 9.4 shows that there is a similar phenomenon with IMF 1-2 of Fig. 9.3

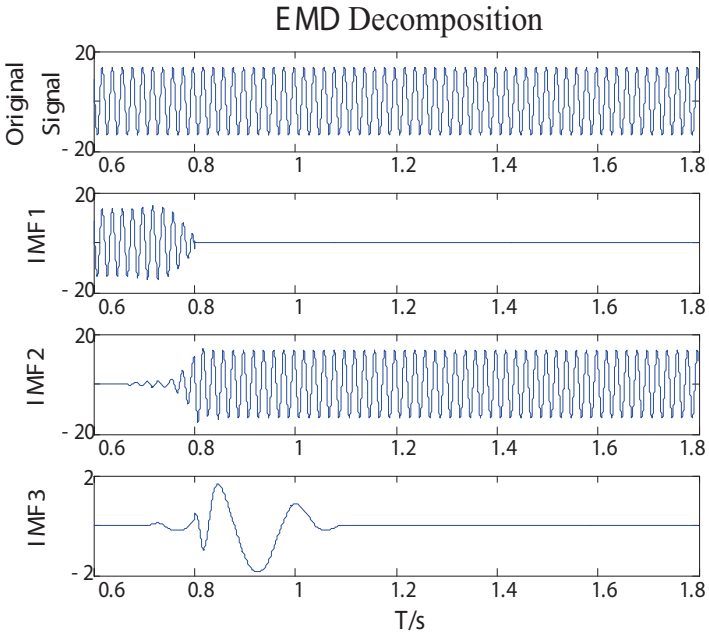


Fig. 9.3 EMD of 2-turn short circuit I_c

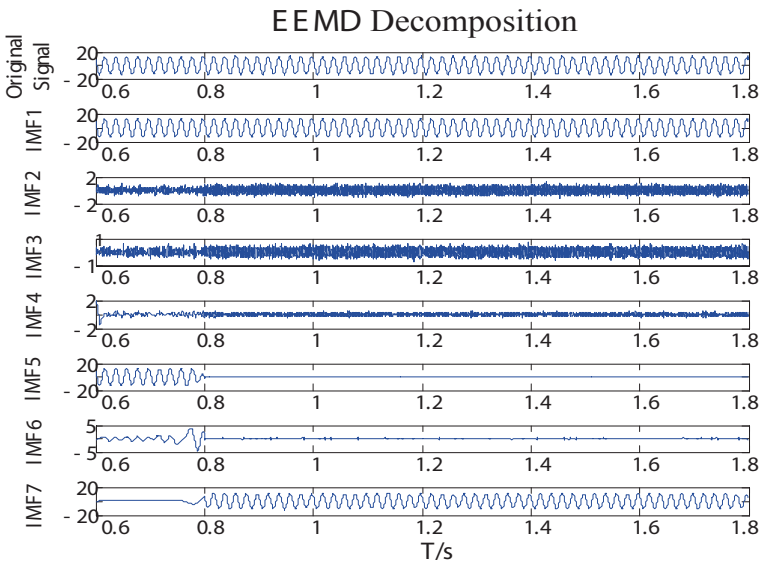


Fig. 9.4 EEMD of 2-turn short circuit I_c

in IMF 5-7. These results prove the feasibility of the two methods in the diagnosis of stator inter-turn short-circuit fault.

Figure 9.3, 9.4 show the analysis which starts from steady state of current; however, when analysis starts from unstable state, some problem will appear as a result of EMD decomposition. When fault occurs at 0.8 s, the analysis result of current I_c of 2-turn short circuit that starts from 0 s is shown as follows:

Figure 9.5 reflects a situation that generator inter-turn short-circuit fault occurs at 0.8 s, and waveform starts at 0 s. Fig. 9.5 shows that there are interference waveforms at fault point as a result of traditional EMD decomposition. Because there are three states (unsteady state, steady state, and fault state) after the generator gets started, the traditional EMD decomposition method cannot effectively distinguish unsteady state from fault state, and the current is divided into three parts. EMD decomposition cannot accurately judge the fault point, which will interfere with the generator fault diagnosis.

But the interference is ruled out after using EEMD. IMF 5-7 in Fig. 9.6 shows that current waveform is accurately divided into two parts—fault-free state and fault state. The method excludes the situation of unstable generator and only retains the characteristics of identifying stator inter-turn short-circuit fault current.

The decomposition results of other situations are similar:

Rotor speed changes from 1200 rpm to 1350 rpm at 0.8 s (Figs. 9.7, 9.8), and 20% drop in stator A-phase voltage happens at 0.8 s (Figs. 9.9, 9.10). IMF 1-2 in Fig. 9.7 shows that the traditional EMD decomposition makes an obvious reaction at the change point 0.8 s. However, there is no obvious reaction in each IMF

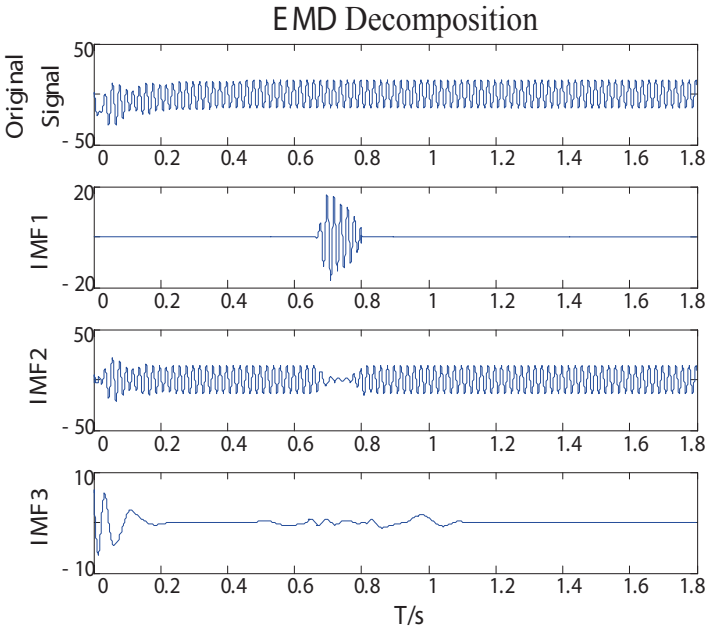


Fig. 9.5 EMD of 2-turn short circuit I_c from 0

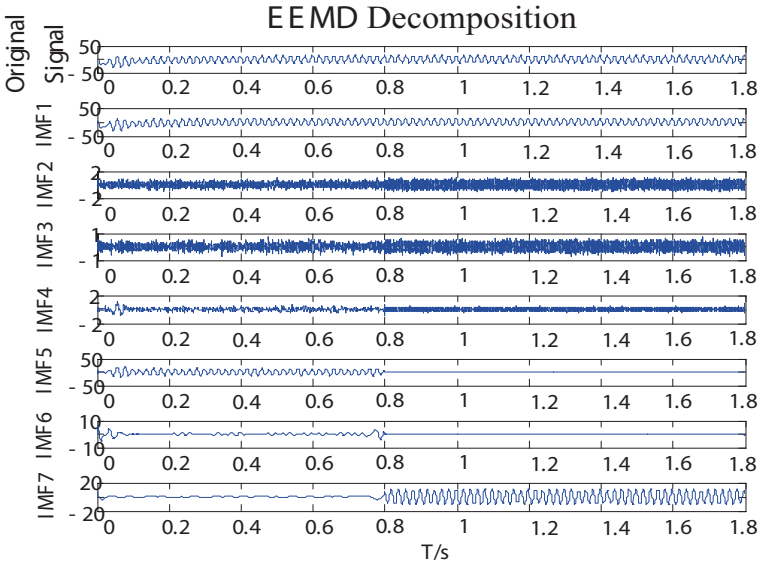


Fig. 9.6 EEMD of 2-turn short circuit I_c from 0

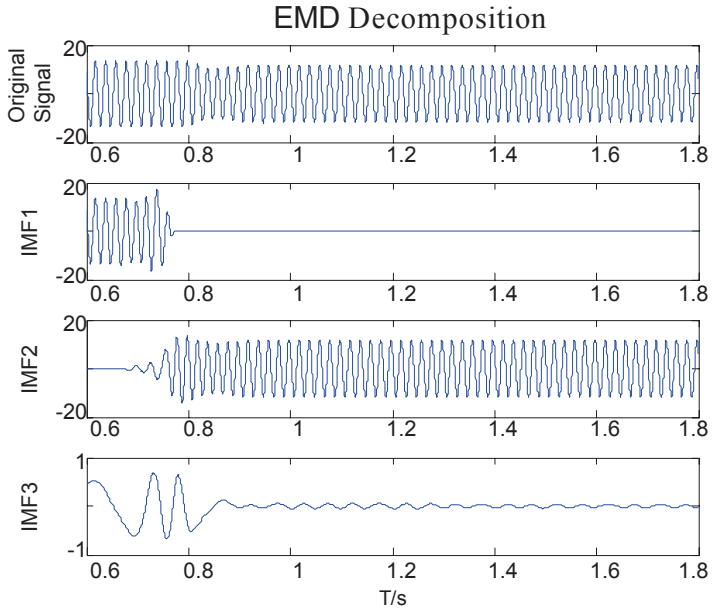


Fig 9.7 EMD of I_c of variable speed

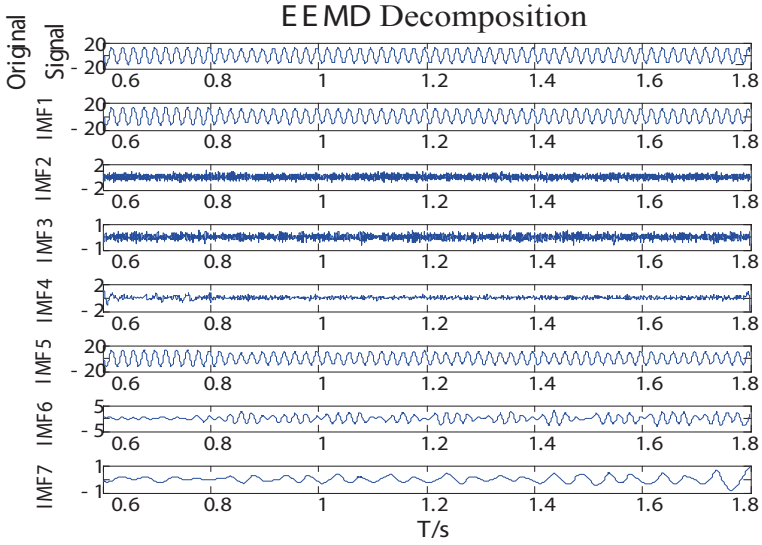


Fig. 9.8 EEMD of I_c of variable speed

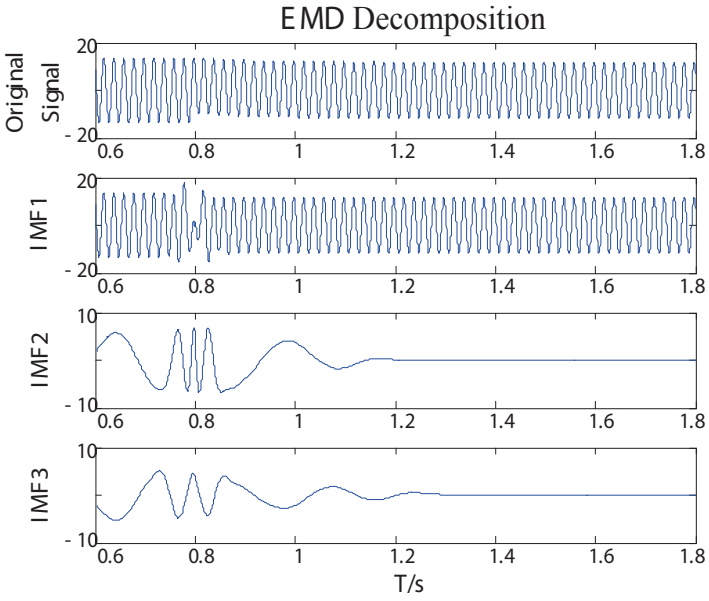


Fig. 9.9 EMD of I_a of voltage drops 20%

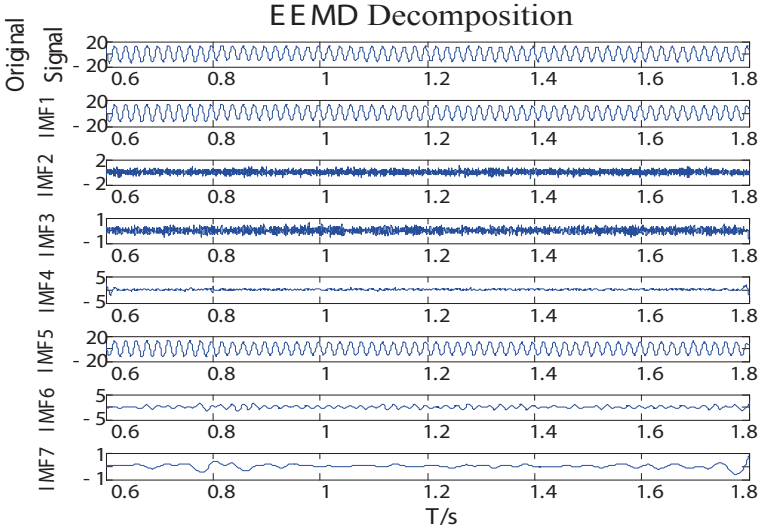


Fig. 9.10 EEMD of I_a of voltage drops 20%

waveform in Fig. 9.8 that is decomposed by EEMD. Change of rotor speed does not make such a decomposition resulting in stator inter-turn short-circuit fault.

Figure 9.9 illustrates that there also is a vibration at 0.8 s point when unbalance fault occurs, although this phenomenon is not similar with others above, the vibration will interfere with the diagnosis process. However, EEMD decomposition does not come up with that result, and there is not obvious vibration at fault point in Fig. 9.10. Therefore, this method does not define voltage unbalance situation as a fault nor makes a better distinction between the two faults.

These results say that EEMD decomposition can reflect the generator fault better and reduce the influence of other factors in fault diagnosis.

9.4.2 Research of Fault Diagnosis

When the generator inter-turn short-circuit fault occurs, harmonic 10, 30, 70, 90, 110, 130, 170 Hz of stator current will increase significantly, and at the same time, harmonic 50, 70, 90, 110, 130, 150 Hz of rotor current will also increase significantly [5].

Harmonic analysis is as shown in Fig. 9.11:

If the inter-turn short-circuit fault is an asymmetric fault inside the generator in terms of macro, then the stator and rotor currents will also contain a harmonic

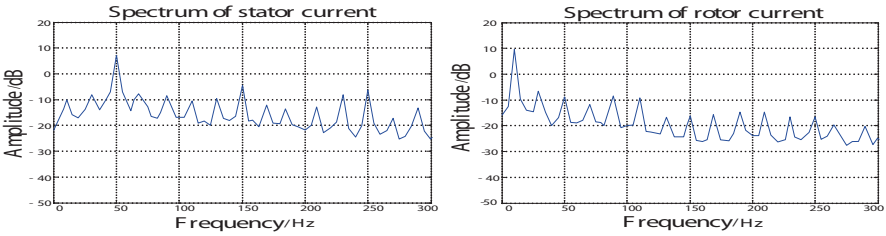


Fig. 9.11 Spectrum of stator and rotor C-phase fault current I_c

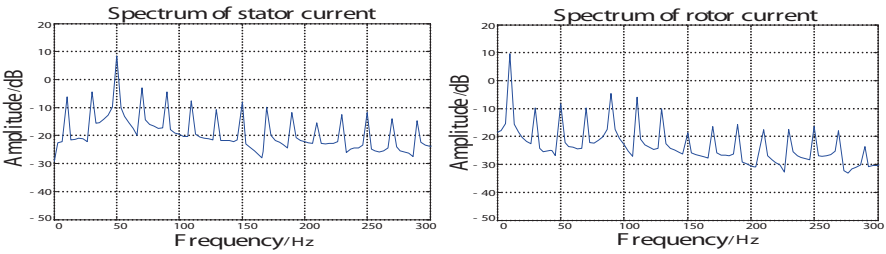


Fig. 9.12 Spectrum of stator and rotor A-phase fault currents I_a

component under the situation of outside voltage asymmetry of the generator. Voltage unbalance fault current I_a harmonic analysis is shown as follows:

As Figure 9.12 shows, when the doubly fed generator stator asymmetric load fault occurs, harmonic increment of stator and rotor currents is similar with inter-turn short-circuit fault; therefore, the two faults cannot be distinguished accurately by judging harmonic analysis merely, and this could cause misjudgment and impact the diagnosis.

There will be harmonic and unbalance current under voltage unbalance situation; therefore, unbalance voltage fault is the most serious interference factor on stator inter-turn short-circuit fault diagnosis. EMD cannot tell the difference between unbalanced voltage and stator inter-turn short-circuit fault; however, EEMD can make stator inter-turn short-circuit waveform change at fault point, while there is no obvious vibration under the situation of voltage unbalance. EEMD method excludes the influence of voltage unbalance when analyzing harmonic.

9.5 Conclusion

Analysis of different situation shows that as the defect of EMD decomposition, there are some wrong judgments in the results of decomposition when the conditions of generator are unstable, rotor speed changes, and voltage is unbalance. However, the defects can be overcome by EEMD decomposition method. By comparing the two methods, EEMD decomposition makes less wrong judgment. EEMD decomposition can not only find out the fault point accurately, but also can effectively exclude interference of other factors on the stator inter-turn short-circuit fault diagnosis. Especially in the analysis process of stator inter-turn short-circuit current harmonic diagnosis, EEMD decomposition can exclude interference of voltage unbalance very well and make the doubly fed induction generator stator inter-turn short-circuit fault diagnosis more reliable.

References

1. Ma H, Zhang Z, Zhang Z, et al. Research on DFIG stator winding inter-turn short circuit fault. Electr Mach Control. 2011;15(11):50–4. (In Chinese)
2. Chen K, Li Y, Chen L. Ensemble empirical mode decomposition for power quality detection applications. Comput Simul. 2010;3(1):263–6. (In Chinese)
3. Zhang X, Wu Y, Zhu S. Bearing fault diagnosis based on ensemble empirical mode and hilbert-huang transform marginal spectrum. Sci Technol Eng. 2010;11(1):7625–9. (In Chinese)
4. Gao J, Wang X, Li F. AC electrical machine and system analysis. 2nd ed. Beijing: Tsinghua University Press; 2005. pp. 1–5. (In Chinese)
5. Li J, Wang D, He L. Steady-state analysis on inter-turn short circuit fault of stator windings in doubly—fed induction generators. Autom Electr Power Syst. 2013;37(18):103–7. (In Chinese)

Chapter 10
Chaotic Property Identification and Prediction of Performance Degradation Time Series for Hydropower Unit

Xueli An

Abstract The performance degradation time series of hydropower unit is reconstructed in phase space by using the chaos theory. Chaotic property of the series is found through analysis. The degradation time series is predicted based on the adding-weight one-rank local-region method. The condition monitoring data of hydropower unit are used to verify the proposed method. The results show that it is feasible to predict the performance degradation of hydropower unit by using the chaos prediction method. The proposed method has high accuracy. It is a new way to operate and maintain the hydropower unit.

Keywords Hydropower unit · Performance degradation · Chaotic property · Adding-weight one-rank local-region method · Degradation prediction

10.1 Introduction

The assessment and prediction of hydropower unit's performance degradation is an important part of the research for the sake of the unit's safe and stable operation. Accurate prediction of performance degradation is the basis for drawing up the unit's maintenance plan. High-accuracy prediction can ensure hydropower units run in more reasonably safe and stable condition. The forecast precision has a direct impact on the hydropower plants' economic benefit; therefore, it is necessary to carry out the performance degradation assessment and prediction to do a lot of research works so as to better ensure the reliability of hydropower units' operation [1].

Due to more factors affecting the hydropower units' performance and nonlinearity of the hydroelectric power system itself, the units' performance degradation time series presents seemingly random characteristics. It is difficult to establish a precise mathematical model to make description; therefore, it is extremely essential to study more reliable and accurate method to predict unit's performance degradation.

X. An (✉)
China Institute of Water Resources and Hydropower Research, 100038 Beijing, China
e-mail: an_xueli@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_10

In this paper, the real monitoring data of a hydropower unit are selected to analyze the chaotic characteristics of this unit's performance degradation time series. Based on this analysis, the degradation time series of the unit's performance has chaotic properties and the phase space reconstruction method is used to establish adding-weight one-rank local-region prediction model [2] for hydropower unit's performance degradation time series. The proposed model is used to predict the changing trend of unit's real degradation time series.

10.2 Chaotic Property Identification of Performance Degradation Time Series for Hydropower Unit

In this paper, the chaos theory is applied to analyze the inherent characteristics of hydropower units' performance degradation, namely, to identify its chaotic characteristics. If the performance degradation time series has chaotic characteristics, the phase space reconstruction theory and the chaos analysis method will be adopted to build the prediction model of hydropower units' performance degradation. The changing regulation of performance degradation is analyzed by using the proposed method.

According to Takens theory, for a time series, when $m \geq 2d+1$ (m is the embedding dimension, and d is the associated dimension of power system), the attractor can be recovered in the m-dimensional reconstructed space, and the phase trajectory of space is reconstructed to remain diffeomorphism for the original dynamic system [3].

When the hydropower unit is in operation, for a given condition parameters time series $\{x(i)\}$, where $i=1, 2, \ldots, n$, if the embedding dimension is m, the time delay is τ, the phase space is reconstructed:

$$Y(i) = [x(i), x(i+\tau), x(i+2\tau), \cdots, x(i+(m-1)\tau)] \tag{10.1}$$

where $i=1, 2, \ldots, N$; $N=n-(m-1)\tau$ as the sampling points of time series. Every phase point contains m components. As to the N phase points, a phase type is constituted in m-dimensional phase space. The change among phase points in m-dimensional phase space describes the system's evolution trajectory in the phase space.

The phase space reconstruction has a pivotal role in chaos prediction. The selection of embedding dimension and time delay are pivotal in phase space reconstruction. If the calculated attractor in the reconstruction phase space has positive maximum Lyapunov exponent, the unit's performance degradation time series has chaotic properties. In this paper, mutual information method [4] is used to compute the delay time. The Cao algorithm [5] is used to compute embedding dimension in phase space reconstruction. The Wolf method [6] is used to calculate the largest Lyapunov exponent of performance degradation time series.

10.3 Prediction of Performance Degradation Time Series for Hydropower Unit

If the hydropower units' performance degradation time series has chaotic characteristics, the chaotic characteristics can be fully utilized to predict the changing trend of performance degradation time series. This can improve the accuracy to a certain extent. In this paper, the adding-weight one-rank local-region method [2] is used to predict the changing trend of hydropower units' performance degradation time series with the specific steps as follows:

Step 1: Look for the nearby points.

In the phase space, the space distance between each neighborhood point and Y_k is calculated, and the reference vector sets Y_{ki} of Y_k are found, where $i = 1, 2, \ldots, q$. The weight of Y_{ki} is defined as follows:

$$P_i = \frac{\mathrm{e}^{-l(d_i - d_m)}}{\sum_{i=1}^{q} \mathrm{e}^{-l(d_i - d_m)}} \tag{10.2}$$

where l is a constant coefficient, generally taking $l = 1$, where d_i is the distance between the points Y_{ki} and Y_k, and d_m is the minimum of d_i.

Step 2: Local linear fitting.

One-order weighted local linear fitting is as follows:

$$Y_{ki+1} = ae + bY_{ki} \tag{10.3}$$

where $e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_m$, $i = 1, 2, \ldots, q$. When the embedding dimension $m = 1$, there is

$$\begin{bmatrix} x_{k1+1} \\ x_{k2+1} \\ \vdots \\ x_{kq+1} \end{bmatrix} = \begin{bmatrix} a + bx_{k1} \\ a + bx_{k2} \\ \vdots \\ a + bx_{kq} \end{bmatrix} \tag{10.4}$$

Based on the weighted least squares method:

$$\sum_{i=1}^{q} P_i (x_{ki+1} - a - bx_{ki})^2 = \min \tag{10.5}$$

The partial derivatives of Eq. (10.5) with respect to independent variables a, b are calculated, respectively:

$$\sum_{i=1}^{q} P_i(x_{ki+1} - a - bx_{ki}) = 0 \tag{10.6}$$

$$\sum_{i=1}^{q} P_i(x_{ki+1} - a - bx_{ki})x_{ki} = 0 \tag{10.7}$$

Solving (10.6), (10.7), a, b are obtained:

$$a = \frac{\sum_{i=1}^{q} P_i x_{ki+1} \sum_{i=1}^{q} P_i x_{ki}^2 - \sum_{i=1}^{q} P_i x_{ki} \sum_{i=1}^{q} P_i x_{ki} x_{ki+1}}{\sum_{i=1}^{q} P_i x_{ki}^2 - (\sum_{i=1}^{q} P_i x_{ki})^2} \tag{10.8}$$

$$b = \frac{\sum_{i=1}^{q} P_i x_{ki} x_{ki+1} - \sum_{i=1}^{q} P_i x_{ki} \sum_{i=1}^{q} P_i x_{ki+1}}{\sum_{i=1}^{q} P_i x_{ki}^2 - (\sum_{i=1}^{q} P_i x_{ki})^2} \tag{10.9}$$

Step 3: Make predictions.

Equations (10.8), (10.9) are substituted into Eq. (10.4). The one-step predicted value $Y_{ki}+1$ of Y_{ki} can be obtained.

Dynamical system is developing and changing with the passage of time. Some disturbance factors of the future will continue to come into the system and will exert influence on the system. Although the history information of system has certain effects in researching it, the predicted value with practical value and higher precision is the most recent data for the prediction time (called the new information). The other farther data only reflect a trend [7].

Therefore, there is no need to use a model to predict all values in the future. In this paper, a known time series is firstly used to establish chaotic prediction model so as to predict a value $Y_{ki}+1$; then, the real value corresponding to the predicted value $Y_{ki}+1$ is added to the known time series with the oldest data point removed meanwhile so that the length of the time series remains unchanged. Afterward, the next value $Y_{ki}+2$ is predicted by using the chaos model. And the latest real value corresponding to $Y_{ki}+2$ is added to the time series, while the oldest data are removed until the forecast is completed.

10.4 Case Study

In this paper, the real condition monitoring data (July 2008 ~ December 2011) of a pumped storage units (rated power 250 MW, rated speed 333 r/min) are used to study the chaotic properties of hydropower units' performance degradation time

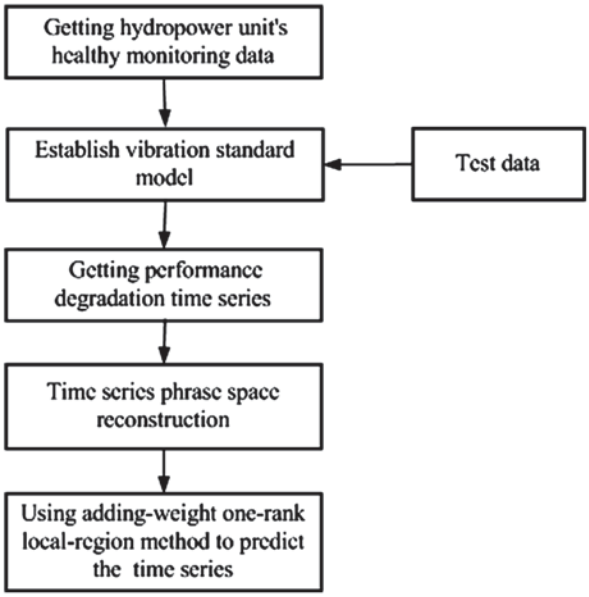


Fig. 10.1 Flowchart of hydropower unit's performance degradation prediction

series and predict them. The vibration is an important indicator of evaluating hydropower units' operational reliability. In this paper, the vibration data of upper bracket horizontal vibration in direction-X are selected as the research subjects. Figure 10.1 demonstrates the flowchart of the hydropower unit's performance degradation prediction.

In order to get a real-time operation condition of the hydropower unit, use the condition monitoring data of unit in the initial running stage and in good without fault condition to establish the vibration standard model of unit in health condition. Considering the important effect of active power and working head on the unit's vibration characteristics, and the good fitting performance of multivariate adaptive regression splines [8] for the scattered data, a vibration-power-working head three-dimensional surface model for hydropower unit is built based on multivariate adaptive regression splines. According to the proposed model, the mapping relationship $v=f(P, H)$ between unit power (P), working head (H), and vibration parameters (v) can be obtained.

As to the 800 sets of data from September 22, 2008 ~ September 18, 2009, 600 sets of data are selected to establish a healthy and standard model with the remaining 200 sets of data used as the test samples for validating the model. In order to make the three-dimensional multivariate adaptive regression splines model have good performance, the selected 800 sets of healthy and standard data should cover the changing range of the working head and the active power work. The active power and working head of 200 sets of test samples are input to the model to compute the healthy and standard values in corresponding condition. The results show

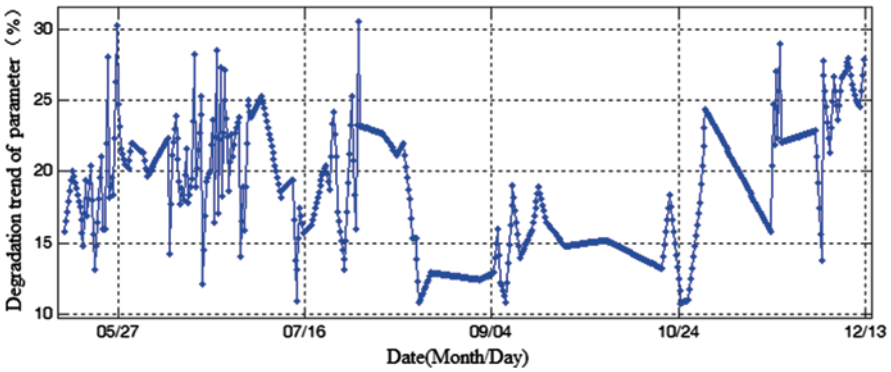


Fig. 10.2 Performance degradation trend graph of hydropower unit

that the healthy and standard values are agreed with the real values. The average relative error is 2.86%.

The real-time online condition monitoring data (the data of active power and working head from May 12, 2011 ~ December 15, 2011) after two years are substituted into the unit health model $v(t)=f(P(t), H(t))$ to get health standard value $v(t)$ in the current operating condition. $v(t)$ is compared with the real value $r(t)$, and the degradation degree $D(t)$ is obtained, where $D(t)=\dfrac{r(t)-v(t)}{v(t)}\times 100\%$, as shown in Fig. 10.2.

It is shown in Fig. 10.2 that the unit's vibration parameter begins to degenerate after two years of operation. The degradation trend has strong volatility and non-stationary. To get the changing trend of the unit performance in the future, it needs to establish a nonlinear prediction model for prediction.

The mutual information method, Cao algorithms, and Wolf method are, respectively, used to calculate the time delay τ, embedding dimension m, and the maximum Lyapunov exponent λ of performance degradation time series (as shown in Fig. 10.2) of hydropower unit. The results are as follows: $\tau=6$, $m=12$, $\lambda=0.1475$. It can be seen that the unit's performance degradation time series has chaotic properties.

According to the calculated time delay τ and embedding dimension m, the phase space reconstruction is made and the adding-weight one-rank local-region method is used to predict the changing trend of unit's performance degradation. The first 593 data are used to reconstruct the phase space; the 594~643 point data are used to forecast the trend. The predicted results are shown in Fig. 10.3.

As shown in Fig. 10.3, the predicted values of unit's performance degradation are agreeable with its real values. The predicted average relative error is 7.52%. The predicted results show that it is feasible to use the proposed method to predict the performance degradation of hydropower unit.

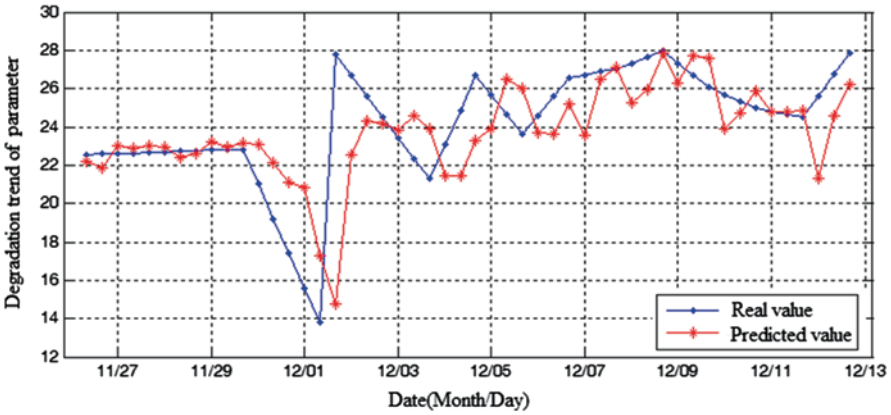


Fig. 10.3 Prediction results of the performance degradation

10.5 Conclusion

The chaos theory is used to analyze and calculate the performance degradation time series of hydropower unit. This study shows that this time series has chaotic property. According to the phase space reconstruction theory, the adding-weight one-rank local-region method is used to predict the changing trend of this time series. The example analysis shows that it is feasible to use the chaos theory to predict the trend of hydropower unit's performance degradation time series. The predicted results are reasonable. The proposed method can timely and accurately estimate the unit's performance conditions and changing trend. It has strong practical value.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Number 51309258) and the Special Foundation for Excellent Young Scientists of China Institute of Water Re-sources and Hydropower Research (Grant Number 1421).

References

1. An XL, Pan LP, Zhang F, et al. Condition degradation assessment and nonlinear prediction of hydropower unit [J]. Power Syst Tech. 2013;37(5):1378–83 (In Chinese).
2. Li S, Zhou Y, Zhu Y. Harmonic current prediction method of combined power filter system based on weighted first-order local area theory [J]. Proc CSEE. 2004;24(6):19–23 (In Chinese).
3. Dong L, Wang L, Gao S. Modeling and analysis of prediction of wind power generation in the large wind farm based on chaotic time series [J]. Trans Ch Electrotech Soc. 2008;23(12):125–9 (In Chinese).
4. Luo H, Liu T, Li X. Chaotic forecasting method of short-term wind speed in wind farm [J]. Power Syst Tech. 2009;33(9):67–71 (In Chinese).
5. Cao L. Practical method for determining the minimum embedding dimension of a scalar time series [J]. Physica D. 1997;110(1–2):43–50 (In Chinese).

6. Wolf A, Swift J, Swinney H. Determining Lyapunov exponents from a time series [J]. Physica D. 1985;16(3):285–317 (In Chinese).
7. An X, Jiang D, Liu C, et al. Wind farm power prediction based on wavelet decomposition and chaotic time series [J]. Expert Syst Appl. 2011;38(9):11280–5.
8. Friedman J. Multivariate adaptive regression splines [J]. Ann Stat. 1991;19(1):1–67 (In Chinese).

Chapter 11
Cosimulation Method for Asynchronous Motor Based on SaberRD–Simulink

Rui Li and Chaonan Tong

Abstract Take the three-phase asynchronous motor as the controlled plant, a comparative study of the cosimulation method is carried out based on the platform which consists of MATLAB/Simulink and Saber. Firstly, build a control model in the Simulink environment, and then compile and set the portable control module by using the tool RTW; call the control module into the SaberRD sketch of the whole asynchronous motor space vector pulse width modulation (SVPWM) control system. Experimental results show that this cosimulation system has the characteristics of flexible application and expected output data, so this study offers a new practical method for research and development of the complex power electronic system.

Keywords SaberRD · Simulink · Cosimulation · Asynchronous motor · SVPWM

11.1 Introduction

The saber platform as power electronic product developed by Synopsys company was born in American, which has been widely used in power electronic, mechanical and electrical integration, mechanical, electronic and optical fields [1]. This software platform features representative current development, advanced level design editing functions, simulation models and analytical skills; support to the standard language and mixed signal modeling; support top-down and bottom-up design method; hybrid analysis technology and wire harness design and advanced cloth line capacity [2].

This paper proposes an effective method for the first time, in combination with the MATLAB/Simulink and SaberRD as united simulation platform for establishing the whole control system [3–4]. It has given the detailed design steps that the rotor flux vector control system for asynchronous motor as the experimental

R. Li (✉) · C. Tong
School of Automation and Electrical and Engineering,
University of Science and Technology Beijing(USTB), 100083 Beijing, China
e-mail: lirui012@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_11

object [5]. Firstly, related researches about the joint simulation of several other schemes based on Windows operating system are introduced; then, in combination with the example, this new joint simulation method is focused based on Simulink/RTW. Experiments show that the method as proposed in this paper runs faster and has higher precision than the traditional single platform. Especially, it is helpful to practical engineering of the real-time development for large-scale power electronic system.

11.2 Comparison of Joint SaberRD–Simulink Simulation Schemes

There are several methods provided in the developed platform with two softwares. These three methods are relatively more effective than others. The first and second method has been used widely for its simple procedure. The third one is used in Saber structure, but the research of application is seldom seen. This strategy is proposed for the purpose of complementing the technique.

1. Cosimulation is used for interactive design tool to connect to the SaberHDL simulator and Simulink. It may include the simulation algorithm of Saber and Simulink, MAST template, and SaberRD simulation module [6]. Because it is a synchronous collaborative simulation, the use of this scheme requires the user to install the MATLAB and the SaberRD on the same computer. As shown in Fig. 11.1, in accordance with the process of the preconfiguration on the software platform [7]:

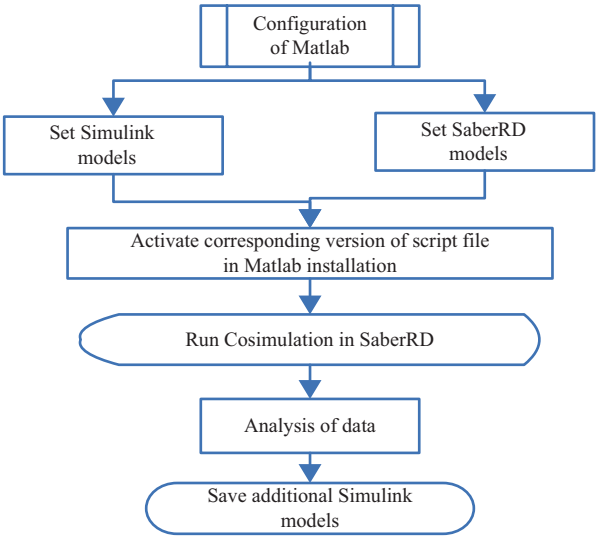


Fig. 11.1 Usage process of cosimulation in SaberRD

Running the simulation with SaberRD as the main body and in the Simulink develop simulation model as SaberRD entity at meanwhile. The key of using this scheme is to combine the simulation as the interface definition, which needs two software environment and reasonable definition for realizing the combined interaction, respectively. In the MATLAB model in the process of development, the Saber Cosim module should be used as the interface function and S-function into Simulink and then save the file as *.mdl [8].

2. The Saberlink function is the transmission of data between MATLAB and Saber interface, the interactive window of special operation in the Saber/CosmosScope page. The data transmission can be dual directional. This method focuses on the data transmission. The model does not involve the interoperability. In order to be suitable to the high real-time request, it does not involve the device and model-invoking circumstances.
3. Use of RTW development control module. This study adopts the scheme of combined simulation. The control module of "portable controller" development process is shown in Fig. 11.2: Five files specified in the configuration file are necessary during the model generation, which have to be found out manually from the SaberRD installation directory and then copied to the MATLAB/Simulink working directory by using the RTW function to call. Among them, "Saber_rtw.tlc" is the control file top target language compiler (compared to the original Simulink RTW grt.tlc file, you can see the MATLAB help file-specific information). "Saber_rtw_foreign.tlc" is in line with the C code generation model of SaberRD programming specifications, including C model database package and desired output. "Saber_rtw_symbol.tlc" is the symbol definition standard SaberRD-based simulation environment, which is used to describe the graphic module from input ports to the output ports. "Saber_rtw_mast.tlc" is based on the MAST language-developed SaberRD template. "Saber_rtw_nt.tmf" plays the role of defining the model of the compile and link interface and generates the dynamic link library CONTROLLERC.dll file.

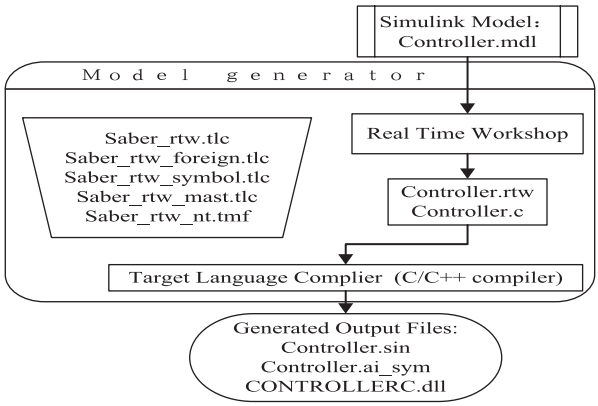


Fig. 11.2 RTW file development process

After the completion of Simulink model development, the control module will be "Build" that needs to do some set of simulation parameters. On the one hand, the simulation type in solver is changed into a fixed step size; for example, a simulation step of SaberRD is t1, the fixed step size in the Simulink simulation sampling time t2, which is generally set to the t1 = t2 or t2 = N*t1. N is a positive integer; on the other hand, run the foreign model in SaberRD with high precision, the simulator will "Warn" promptly, and the results may be distorted seriously. With the steps to configure RTW, specify the system target file "saber_rtw.tlc," and the appropriate user variable is set in the computer, thus pointing to the C program compiler.

11.3 Vector Control System of the Induction Motor

The space vector pulse width modulation technology is used in combination with the simulation method described in this paper, the rotor flux of induction motor indirect oriented control. The control system is called the vector control (VC) system or the rotor flux-oriented control (FOC) referred system by the orthogonal coordinate system, equivalent to DC motor model [9]. The application of the double closed loop controls the structure of the system, in which the outer-loop feedback is the speed of the motor and feedback loop is the torque current and exciting current [10]. The structure diagram of control system is shown in Fig. 11.3.

The stator terminal current is shown in Fig. 11.3, in which the three-phase asynchronous motor can be measured by instrument, and then through the Clark coordinate transformation, the rotating magnetic field (a, b, c) three-phase coordinates model into a model two-phase synchronous rotating reference coordinates [11]. Then (α,β) current vector in the coordinate system through Park transform into the rotor flux orientation rotation (d, q) coordinate model. Selection of state variables $\mathbf{X}=\begin{bmatrix}\omega & \psi_{rd} & \psi_{rq} & i_{sd} & i_{sq}\end{bmatrix}^T$, including "d," represents stator current flux direction, "q" represents torque direction, "represent s" stator "r" represents stator and rotor. ω is the motor output speed. The input variable is $\mathbf{U}=\begin{bmatrix}u_{sd} & u_{sq} & \omega_1 & T_L\end{bmatrix}^T$, ω_1 is the synchronous speed of the system, the output variable is $\mathbf{Y}=\begin{bmatrix}\omega & \psi_r\end{bmatrix}^T$ while considering squirrel cage asynchronous motor rotor is short, so we have $u_{rd}=u_{rq}=0$. The torque equation system is shown as follows:

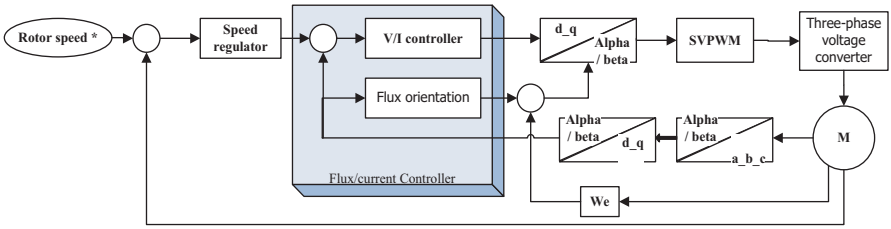


Fig. 11.3 Diagram of double closed-loop control system of asynchronous motor

$$T_e = \frac{n_p L_m}{L_r}(i_{sq}\psi_{rd} - i_{sd}\psi_{rq}) \tag{11.1}$$

The asynchronous motor state equation is shown in the formula below (Eq. 11.2):

$$\begin{cases} \dfrac{d\omega}{dt} = \dfrac{n_p^2 L_m}{JL_r}(i_{sq}\psi_{rd} - i_{sd}\psi_{rq}) - \dfrac{n_p}{J}T_L \\ \dfrac{d\psi_{rd}}{dt} = -\dfrac{1}{T_r}\psi_{rd} + (\omega_1 - \omega)\psi_{rq} + \dfrac{L_m}{T_r}i_{sd} \\ \dfrac{d\psi_{rq}}{dt} = -\dfrac{1}{T_r}\psi_{rq} - (\omega_1 - \omega)\psi_{rd} + \dfrac{L_m}{T_r}i_{sq} \\ \dfrac{di_{sd}}{dt} = \dfrac{L_m}{\sigma L_s L_r T_r}\psi_{rd} + \dfrac{L_m}{\sigma L_s L_r}\omega\psi_{rq} - \dfrac{R_s L_r^2 + R_r L_m^2}{\sigma L_s L_r^2}i_{sd} + \omega_1 i_{sq} + \dfrac{u_{sd}}{\sigma L_s} \\ \dfrac{di_{sq}}{dt} = \dfrac{L_m}{\sigma L_s L_r T_r}\psi_{rq} - \dfrac{L_m}{\sigma L_s L_r}\omega\psi_{rd} - \dfrac{R_s L_r^2 + R_r L_m^2}{\sigma L_s L_r^2}i_{sq} - \omega_1 i_{sd} + \dfrac{u_{sq}}{\sigma L_s} \end{cases} \tag{11.2}$$

where R, L, ψ, I, u represent the resistance, inductance, flux linkage, and the state variables of current/voltage, respectively. L_m is the mutual inductance. Subscripting characters "d" and "q" note two phase rotating coordinates as d-axis and q-axis. And the "d," "r" express on behalf of stator and rotor, respectively. J is the inertia of the motor, n_p is the number of pole pairs, ω is the angular velocity of the mechanical rotor, and ω_1 is the angular velocity of d_q coordinate system. Among them, the electromagnetic time constant of the rotor is $T_r = \dfrac{L_r}{R_r}$, and the motor's magnetic flux leakage coefficient is $\sigma = 1 - \dfrac{L_m^2}{L_s L_r}$.

In the control loop of double closed-loop feedback, the (d, q) two-phase flow of power decoupling is individually controlled by the PI controller. The speed controller is built in MATLAB/Simulink. The desired speed of idse_cmd by the input signal input, and the rotating shaft of the motor feedback speed signal idse compared. The difference is used as a speed controller input control signal. A closed loop feedback, decoupling current iqse current iqse* will generate torque as the input and the asynchronous motor external feedback are compared, the difference as the input signal through the current/voltage controller (developed in MATLAB/Simulink).

The slip output calculation module and the rotor speed are equal to the value of the superposition of synchronous speed of ω_1. The output is equal to the magnetic field of the rotor slip angle relative to the stator magnetic field. Driving circuit of asynchronous motor is developed in SaberRD. Its role is to control voltage vector pulse width modulation module using PWM modulation and the SVPWM three-phase bridge drive circuit for the power supply of three-phase asynchronous motor [12]. The drive circuit is composed of six switch devices. Different combinations of

the three bits binary number according to the sequence of A, B, and C represent different states of the switch, in which "1" means the upper bridge arm power device turn on and the lower bridge arm device to turn off; "0" represents that the upper bridge arm power device is closed and the lower bridge arm device to turn on. Each state corresponds to a SVPWM voltage waveform a pulse wave [13].

According to different applications, select three or more than three vectors to calculate the synthetic vector. According to the parallelogram law:

$$T_S V = T_1 V_1 + T_2 V_2 \tag{11.3}$$

where T_1, T_2 into the vector V_1, V in the switching cycle duration; T_S PWM switch cycle. Let T_0 be a continuous time, and zero vector of V_0 is $T_S = T_0 + T_1 + T_2$. If the angle between the V and V_1 is θ, $T_1 = mT_S \sin(\pi/3 - \theta)$, and $T_2 = mT_S \sin(\theta)$, where m is SVPWM modulation coefficient, $m = \frac{\sqrt{3}}{V_{dc}} \left| V_{ref} \right|$.

To sum up, the system simulation master plan development cosimulation method (including the control module, driver circuit and asynchronous motor and the mechanical load). The controller "ASR" contains three control modules, which are generated by the RTW method.

11.4 Analysis of Simulation Results

Test specification: Operating system: Windows xp or Windows7. Software requirements: SaberRD_2012 version, MATLAB_2010b or higher version, the VC++_6.0 compiler. Operation simulation platform and the signal for partially are shown in Table 11.1:

Table 11.1 List of experimental variables

Signal meaning	Variables (unit)
Signal output module PWM	g1~g6 (-)
Input voltage	va, vb, vc (V)
The stator terminal current	ias, ibs, ics (A)
d-axis reference input current	ids_cmd, idse (A)
q-axis reference input current	iqs_cmd, iqse (A)
Speed input	pd_cmd (rad/min)
Slip regulating reference value	Swe (-)
Motor slip measurements	slip_percent (-)
Motor speed measurement	w_rpm (rad/min)

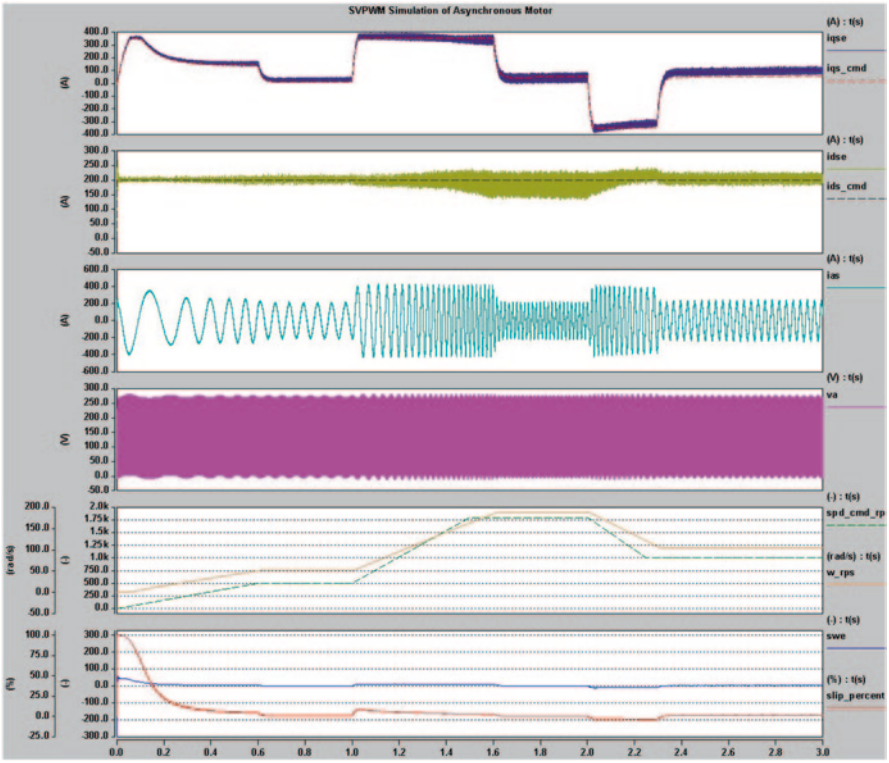


Fig. 11.4 Cosimulation of VC system for asynchronous motor in SaberRD

Set the parameters of simulation environment: the simulation length: 5 s; MATLAB model simulation step: 0.001s; SaberRD simulation step: 1e-5 s; integral order number: 1; target iteration number: 10.

The control module parameters: conversion rate: slew_rate = 200; speed controller: kp = 1, ki = 1; voltage/current controller: kpq = kpd = 0.1, kiq = kid = 10.

Parameters of asynchronous motor: Lm = 0.001H: Ls = Lr = 0.2e-3H; mutual inductance; stator resistance: Rs = 0.01 Ω; the equivalent rotor resistance: rr = 0.006 Ω; motor pole pair number: p = 4; moment of inertia: J_Motor = 0.04 kg.m^2.

SVPWM module parameters: reference voltage: V_rail = 266V; sinusoidal carrier frequency: 5000 Hz.

The simulation results are shown in Fig. 11.4. The results show that the input command speed: from zero to a slow ramp, with greater slope of the ascending velocity, to run at full speed, then slow down to normal speed level test. Small mechanical load for testing is 1kg/ m^2. The key of slip in the process of motor running is the precondition of rotor rotation and torque. Comparing the measured slip value and reference value, we can enter the tracking control of steady state when the motor is started after 0.2s through the slip calculation module of dynamic adjustment. But in the set, the measured values are still biased. By changing the controller

appropriately or regulating the input flux calculation model to achieve more precise control, as shown in Fig. 11.4, the torque and rotor flux-oriented control will have good effect. For both speedup and speed-down processes, for the current component iqse, torque can be synchronously tracking the set value iqs_cmd. When the flux variation features better coordination and speed setting, its volatility is higher with better control effect by adjusting the parameters of the speed controller. Observation and analysis of driving circuit for high-frequency carrier in SVPWM work.

With each state is symmetrical, the six-switch states in circuit SVPWM and the related voltage and current were analyzed. The results show that some of the high switching frequency 5 kHz, about 2.5 kHz which is enough to drive motor to achieve better control effect. Of course, if low-frequency control signal is applied, greater harmonic distortion will be caused.

11.5 Conclusion

A new cosimulation method is proposed in this paper. Respectively, the two parts of one system are developed in SaberRD and Simulink platform, which not only makes full use of advantages of Simulink in analyzing the mixed signal, but also makes the control algorithm embedded in the standard library of Saber. By calling the RTW tool, a portable control module is generated. Finally, in the SaberRD environment, a complete SVPWM control system is built. The simulation results show that the system is built by using the cosimulation method which is accurate and good and shows dynamic, static performance, especially and much faster than traditional method. Considering the electrical engineering in the separate and independent control and design requirements, this paper provides an effective tool for the design and development for complex power electronic system.

Acknowledgements This paper is supported by the Research Project of State Key Laboratory of Mechanical System and Vibration (MSV201409) and the Beijing Fundamental Research Funds for the Central Universities (FRF-SD-12-008B).

References

1. He X, He J. Simulation study of exciter of AC excitation generator based on Saber. Power Sys Prot Control. 2009;37(10):90–2.
2. XUE G, PAN Q, WU LI Wen-. Research on systematic simulation model of 3-phase 3-level inverter. Electr Mach Control. 2014;18(4):18–23.
3. Hu YZ, Guo Su Na. Asynchronous motor vector control system based on space vector pulse width modulation, mechatronics and automatic control systems. Lect Notes Electr Eng. 2014;23(7):675–82.
4. Ma H. PMSM sensorless vector control system based on single shunt current sensing, mechatronics and automatic control systems. Lect Notes Electr Eng. 2014;23(7):727–35.

5. Zeng W, Wang J. Saber and simulink cosimulation of vehicle mounted inverse power system. Mod Electron Technol. 2012;35(12):186–8.
6. Synopsys Corporation. MATLAB/Simulink to Saber model export usermanual. US: synopsyscorporation. Version1; 2011. p. 138–141.
7. Vanijjirattikhan R, Ayhan B, Tipsuwan Y, et al. A web based distributed dynamics simulation of a three-phase induction motor. Proceedings of NSF Workshop on Teaching of First Courses on Power Electronics and Electric Drives and Advance Courseon Power System Applications of Power Electronics; Tempe, Arizona; 2003. p. 10–15.
8. Wang P, LI F, Zhao J. Cosimulition of saber and simulink –based ppt satellite power controller. J Austronautics. 2010;31(12):2819–22.
9. LIU T, TAN Y, WU G. Simulating of high speed PMSM control system based on SVPWM [J]. Power Syst Prot Control. 2009;37(12):11–4.
10. ZHANG Y, ZHAO Z, LU T. Control system for three-level dual-PWM converter based on RTW. Electr Power Autom Equip. 2009;29(10):128–31.
11. Mastorocostas C, Kioskeridis I, Margaris N. Thermal and slip effects on rotor time constant invector controlled induction motor drives. Power Elect IEEE Trans. 2006;21(2):495–504.
12. Gao H, Chen H, Wu M. Simulation of voltage space vector PWM method for AC motor. Marine Electr Electr Eng. 2007;27(3):129–33.
13. LI R, Zheng X, BAI Y. Simulation of SVPWM inverter control based on saber. Mod Electr Technol. 2011;34(16):203–6.

Chapter 12
Research and Development of Prejudging Device for Open Circuit in Secondary Side of Current Transformer

Gaoxiang Lin, Bifu Qian, Rongtao Wang, Xijia Liu and Jiaxing Wang

Abstract The safe and reliable prejudging device for open circuit in the secondary side of current transformer has been developed. On the basis of explaining the process and defects of the traditional operation, the paper designs a schematic diagram of prejudging device. The suitable values of the shunt resistors have been obtained from lots of calculation and analysis with the safety precaution and approach designed at meanwhile including the components selection and the design optimization. The test results show that the device could well prejudge the circuit open in the in-mending section of the circuit at the secondary side of the current transformer.

Keywords Current transformer · Open circuit · Shunt resistor · Prejudging

12.1 Introduction

The current transformer (hereinafter referred to as CT) is an important part of the protection subsystem or the measure & control subsystem in the power system. Its operating stability may affect the accuracy of the measure & control subsystem as well as the reliability of the protection subsystem [1–2].

The current in the primary winding of CT is called as the primary current which is represented by I1. As the same time, the current in the secondary winding of CT is called as the secondary current which is represented by I2. The ratio of I1 and I2 are defined as the CT ratio which is represented as K:

$$K = I1/I2 \tag{12.1}$$

The schematics of CT were shown in Fig. 12.1.

G. Lin (✉) · B. Qian · R. Wang · X. Liu · J. Wang
Wenzhou Power Supply Company, 325000 Wenzhou, China
e-mail: lingaoxiang@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_12

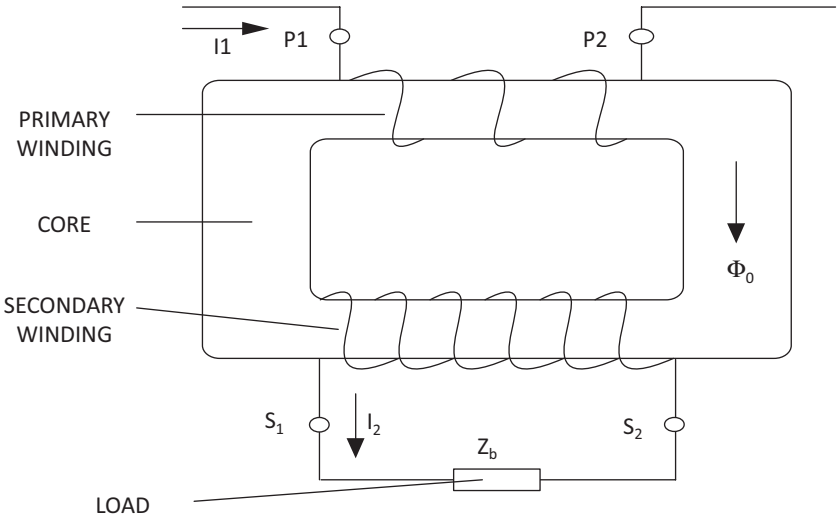


Fig. 12.1 Schematics of CT

If the circuit in the secondary side of CT is open, the primary current will change to the excitation current totally, which will thus cause the iron core of the CT excessively saturated to produce high voltage while the peak value will be available at several kilo-voltages between S1 and S2; therefore, the open circuit in secondary side of CT is harmful [3–5].

"When workers are working on the circuit in the secondary side of a working CT, it is forbidden to open the secondary side circuit", quoted from the clause 10.13 of *The Safety Regulations for Electric Power* which was released by the General State Administration of Quality Supervision, Inspection and Quarantine, and National Standardization Management Committee of the People's Republic of China.

There was no effective method to prevent the in-mending section of the circuit open in the secondary side of a working CT. The method for dealing with the problem was mounting an over-voltage protection [6–7], which could reduce the range and degree of accident, but could not avoid it.

The main content and achievements of this article are shown as below:

Firstly, the paper has elaborated the traditional protection method of in-mending section of the circuit open in the secondary side of a working CT, and informed the defects.

Secondly, on the basis of lots of statistics and calculation, this article has proposed a prejudging device for open circuit in secondary side of CT based on the principle of shunt resistors, which could work well in practice. Also, its safety measures are listed.

Thirdly, the experimental results show that the device could prejudge the open circuit in secondary side of CT correctly, and achieve the desired effect.

12.2 Working Principles of the Device

There are two positions, as shown in Fig. 12.2 for mounting the shoring stub. The first position is POSITION A, where the shoring stub is mounted in the case of overhaul or reforming for the device A; the second position is POSITION B for the same working for device B.

Whether the operation of mounting the shoring stub is successful or not can be judged by observing the current values on the terminal device (e.g. DEVICE B in Fig. 12.2). However, there is no reliable method for the operation of resuming of making sure the integrity of the current circuit. The traditional operation of resuming is described in details as below, taking the example of reforming the DEVICE B shown in Fig. 12.2.

STEP 1: check the integrity of the current circuit of the DEVICE B and the terminal block of the CABINET B.

STEP 2: check the reliability of the current connecting pieces in the terminal block of the CABINET B.

STEP 3: one worker removes the shoring stub on the POSITION A phase by phase; another one keeps watching the current values.

There are three unstable factors for resuming the traditional operation by analyzing the process above:

FACTOR 1: how to ensure the integrity of current circuit between the terminal block and the DEVICE B (including the DEVICE B itself) before the current connecting pieces are removed.

FACTOR 2: how to ensure the reliability of the connection of the current circuits between the terminal block and the DEVICE B (including the DEVICE B itself) before the current connecting pieces are removed. The resistance value of the circuit in secondary side of CT will be larger than normal when the screws have not been screwed tightly, or the connecting pieces have not been connected reliably, or in case of some similar situations. It is difficult to find out the defects, especially when the fault happens in N phase.

FACTOR 3: how to ensure the reliability of the connect pieces.

The device's structure is shown in Fig. 12.3. The resistance value between phase A (or B, or C) and phase N can be controlled by setting the three-location toggle switch K. when toggle switch K is in Location 2, the resistance value between

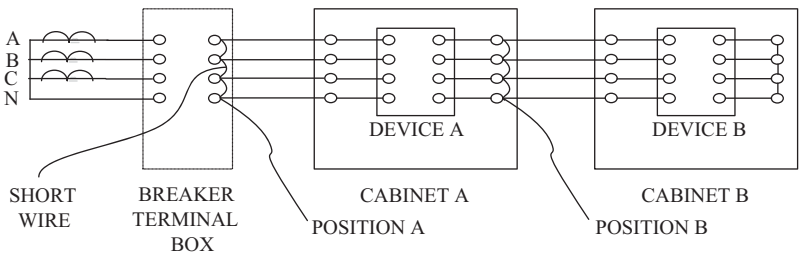


Fig. 12.2 Diagram of the circuit in the secondary side of CT

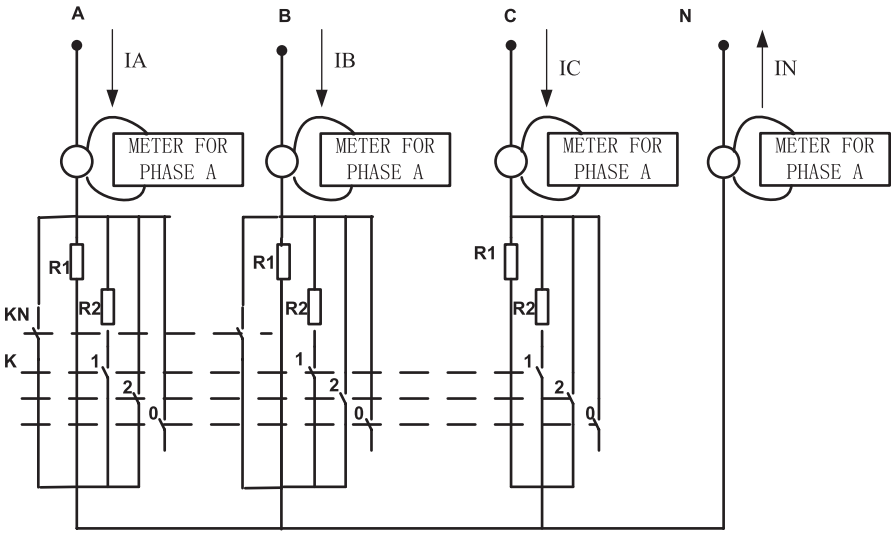


Fig. 12.3 Diagram of the device's structure

phase A (or B, or C) and phase N will be zero, equivalent to the short-circuit; when in Location 1, the resistance value is R1||R2, which is defined as R12; when in Location 0, the resistance value is R1. The resistance value of R1 is approximately equal to the resistance value of transmission line between breaker terminal box and the DEVICE B, including the internal resistance value of the DEVICE A. Also the same, the resistance value of R12 is approximately equal to the resistance value of cable between the DEVICE A and the DEVICE B, including the internal resistance value of the DEVICE B.

If the resistance value of the load in the secondary side of CT is too large, the CT will work in nonlinear and cause wrong actions of the protection device and wrong samples of measure & control device; therefore, the resistance values of R1 and R2 have to be set appropriately: a large number of transmission lines of 110 kV substations have been inquired for their model and length. Upon calculation and analysis, we've finally gotten the values of R1 and R2, respectively, 0.5 and 0.2 Ω.

The method of prejudging circuit open in phase A (or B, or C), is observing the current of phase A (or B, or C) changing via switching the three-location toggle switch K. The button switch KN is set for prejudging circuit open in phase N.

12.3 Development of the Device

12.3.1 Selection and Design of Device

1. The power source—laminated battery, which is small and has a high output voltage and excellent performance. It is suitable for portable instrumentation.

Fig. 12.4 Diagram of sampling program of MPS430

```
uchar dis3[6]
uchar codedis4[]={"080459122803"};
void delay_1ms(uint x)  //delay
{
    uint i,j;
    for(j=0;j<x;j++)
        for(i=0;i<110;i++);
}
void AD()   //A/D convertion module
{
    uchar adval, i;
    for(i=0;i<10;i++)
    {
        dianya = 0;
        csad = 0;   //cs start
        adwr = 1;
        _nop_();
        adwr = 0;   //A/D convertion starts
        _nop_();
        adwr = 1;
        while(intrl = 0); //waiting for converting finished
        AD_DATA = 0xff;     //anti-inference
        adrd = 1;
        _nop_();
        adrd = 0;   //read the data
    }
}
```

2. The regulator circuit—78L05 integrated voltage regulator chip, which is small and suitable, and has an output of 6–20 V DC.
3. The regulating switch of resistance value—three-location toggle switch.
4. The processor—MPS430 microcontroller, although has a higher price than STC89C51, a lower power consumption and a larger memory. Its sampling program is shown in Fig. 12.4 with its hardware wiring diagram shown in Fig. 12.5.
5. The current sensor—TLGK Rogowski coil has a good impact resistance—more than 100 A and a high accuracy class—0.1%. Its connection characteristic is direct-connection [8–10].

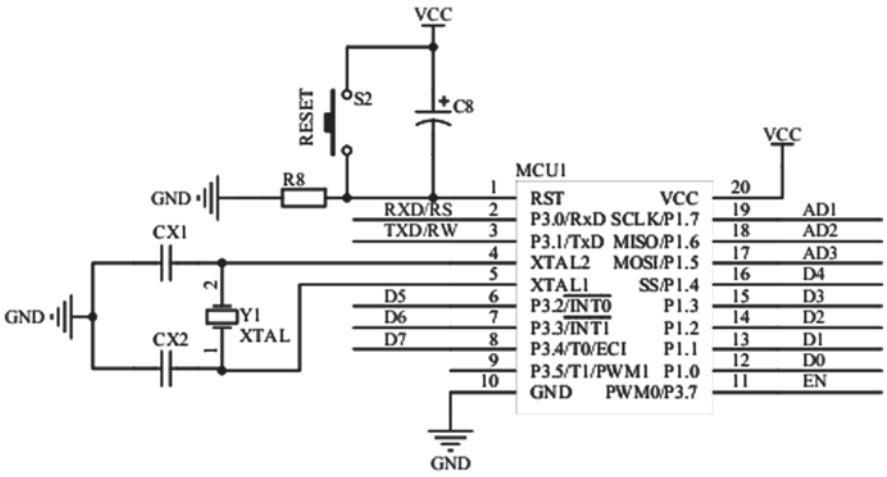


Fig. 12.5 Diagram of hardware wiring of MPS430

```
Static struct fb_ops s3c2410fb_ops =
{
    owner = THIS_MODULE;
    fb_set_par = s3c2410fb_check_var;   //detect parameters
    fb_set_par + s3c2410fb_set_par;     //set display pattern
    fb_blank = s3c2410fb_blank;         //turn on or turn off
    fb_setcolreg = s3c2410fb_setcolreg; //set the registors' values
    fb_copyarea = cfb_fillrect;         //select a rectangular region
    fb_copyarea = cfb_copyarea;         //copy and display
    fb_imageblit = cfb_imageblit;       //display
};
```

Fig. 12.6 Diagram of data-processing program of the LCD display

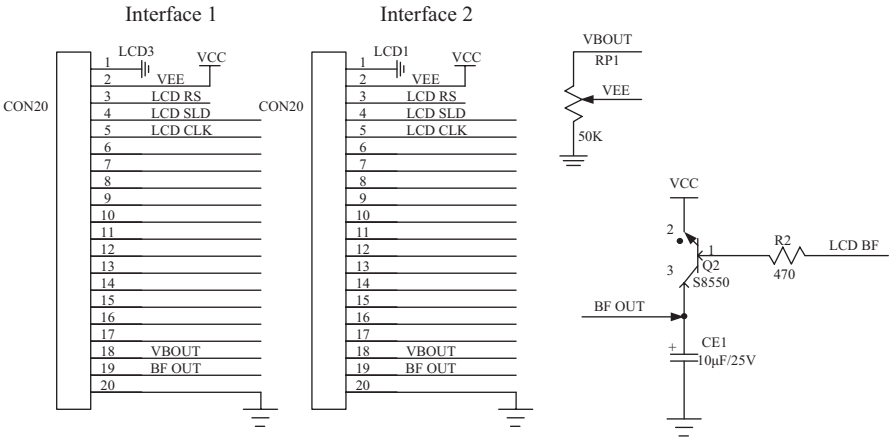


Fig. 12.7 Diagram of hardware wiring of the LCD display

6. The data display module—LCD display. Its data-processing program is shown in Fig. 12.6 with its hardware wiring diagram shown in Fig. 12.7.

7. The shunt resistor—RX24G aluminum-shell resistor, whose rated current and shock current meet the requirements and have a high accuracy class.
8. The shell of the device—engineering plastic insulation.

Upon assembling and installing the eight modules above, we've gotten the device which is shown in Fig. 12.8.

12.3.2 Safety of the Device

Firstly, 0.2 and 0.5 Ω are adopted as the resistance value of R12 and R1 respectively. The meaning of these two values indicate the resistance value of the load in secondary side of CT should be never larger than the rating secondary resistance value when the prejudging device is in parallel with the in-mending device, or the prejudging device is in series into the secondary current circuit

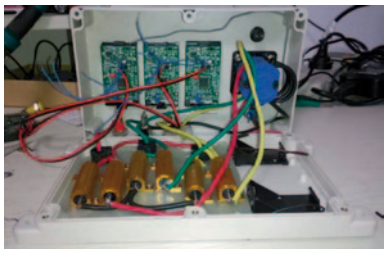


Fig. 12.8 Diagram of the interior of the device

Secondly, the resistor R1 is welded inside the device directly to avoid the short-time open in the process of switching the toggle switch K.

Thirdly, the toggle switch K, the Rogowski coil and the aluminum-shell resistor we select could withstand the impact of 200 A current lasting 100 ms at least; therefore, the components may work as normal when a short-circuit fault happens.

Why 200 A? The 150 A is the largest current of the CT of 10P30 accuracy class in linear region, so 200 A is even larger than 150 A.

And why 100 ms? The inherent time of the breaker acts in a short-circuit fault was about 60–80 ms, so 100 ms even is larger than 80 ms.

As a result, the device can work properly in the most serious fault with the parameters of 200 A and 100 ms.

Fourthly, the measurement part and the realization part of the device are connected by the electromagnetic coupling of the TLGK Rogowski coil to avoid the interference in the measurement part.

Fifthly, when the device is mounted, the toggle switch K should be set in Location 2(short-circuit) for reducing the adverse impact because the resistor R1 or R2 would emit heat when it keeps working for a long time.

12.4 Function Test

12.4.1 Method of Device Use

Generally, the device in the secondary side of the CT, such as the protection device, the measure & control device and the BZT device, has tripping function depending on accurate data; therefore, the prejudging device should have strict operating steps. The paper took Fig. 12.2 as an example (the DEVICE B needed mending) to explain the steps.

Firstly, mount the prejudging device phase by phase at the Position B, as shown in Fig. 12.2 after switching the toggle switch K to Location 1.

Secondly, check the current values of the three phases and made sure that they are balanced and changed in the same way when switching the toggle switch K; nevertheless, the device has some problems and needs overhaul.

Thirdly, disconnect the circuit between the CARBINET A and the CARBINET B, at Position B; then overhaul DEVICE B.

Fourthly, connect the circuit between the CARBINET A and the CARBINET B at the Position B after overhauling DEVICE B.

Fifthly, switching the toggle switch K to Location 2 or Location 0 can check the changing of the current values. If the current values of one phase (or two phases, or even three phases) keep still, the corresponding current circuit of that phase (or those two phases, or even those three phases) is open. If the current values change inconsistently, there are some unreliable connections in some phases.

Sixthly, after making sure the integrity of the current circuits in phase A, B and C, switch the toggle switch K to Location 1 (short-circuit), and record the values of IA, IB, and IC; then switch K to Location 2, press the button switch KN, and observe the changing of current value in phase N. If IN is almost equal to 10%*IA (IA=IB=IC, generally), the current circuit is integral; otherwise, the current circuit in phase N is open.

Seventhly, remove the test wires of the prejudging device.

12.4.2 Test Results

The prejudging device has been tested on the simulation test platform for the correctness of the principle as we proposed and the reliability of the device. The simulating parameters are set as below.

Firstly, simulation of DEVICE A as shown in Fig. 12.2 is in overhauled condition, and the prejudging device is connected at Position B in a parallel manner;

Secondly, the current value from the secondary side of CT is set as 2 A.

Thirdly, the resistance value between the DEVICE A and DEVICE B (including the DEVICE B itself) is set as 0.5 Ω and the resistance value of the Prejudging Device is set as 0.5 Ω.

The test results are shown as below.

When the current circuit between the DEVICE A and DEVICE B is an integral, as shown in Fig. 12.9b due to the same resistance value of the two branches (the one branch is the circuit between the DEVICE A and DEVICE B including the DEVICE B itself, and another one is Prejudging Device), the current values shown on device are balanced, as shown in Fig. 12.9a.

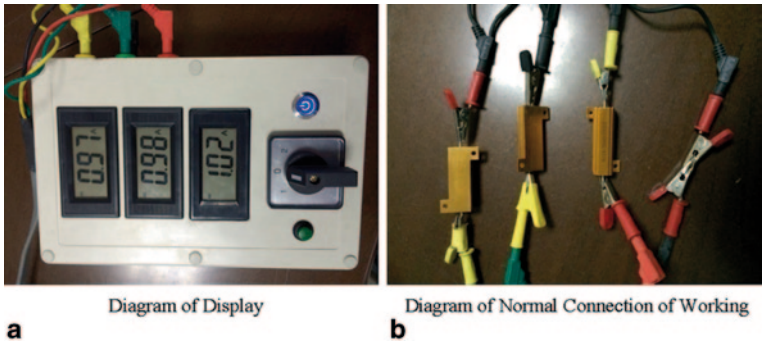


Fig. 12.9 Diagram of normal simulation

Fig. 12.10 Diagram of simulating open circuit

When the current circuit between the DEVICE A and the DEVICE B is open at A phase, as shown in Fig. 12.10b. Obviously, the current can't go into the circuit between the DEVICE A and the DEVICE B in Phase A at all, so the prejudging device would bear 2 A in Phase A. Because of Phase B and Phase C are complete, the currents of the two phases are balanced too. The current values on device were shown as Fig. 12.10a.

12.5 Conclusion

A safe and reliable prejudging device based on the principle of shunt resistors for open circuit in the secondary side of CT has been developed. It can ensure the integrity of the in-mending section of the circuit in the secondary side of CT when the CT in working. The test results show that the device could prejudge the open circuit in the secondary side of CT correctly and achieve the desired effect.

References

1. Hongquan J, Jian Z, Yansong Y, et al. Comment on performance of differential protection taking into account electronic current transducers. Power Syst Technol, Guangzhou. 2006;30(23):61–6, (in Chinese).
2. Yubo Y, Yuping L, Yang X, et al. The influence and countermeasure to transformer differential protection of CT partial transient saturation caused by external fault removal. Proc CSEE, Guangzhou. 2005;25(10):12–7, (in Chinese).
3. Luciano BA, De Albuquerque JMC, De Castro WB, et al. Nanocrystalline material in toroidal cores for CT analytical study and computational simulations. Mater Res, Washington. 2005;8(4):395–400.
4. Wei X, Hang L, Li L, et al. Distinguish CT's extreme saturation by waveform discrimination of branch current. Electr Power Autom Equip, Nanjing. 2005;25(9):39–42, (in Chinese).
5. Yong L, Youyuan N. 3-D magnetic field analysis and mutual inductance computation of current transducer. Electr Power Autom Equip, Nanjing. 2011;31(1):27–30, (in Chinese).

6. Guoqing C. The influence of saturation of the CT on relay protection. Tech Autom Appl, Harbin. 2007;26(10):115–6, (in Chinese).
7. Ruisheng L, Guang-hui L, Qiang W. A Saturation criterion of CT for line differential protection. Electr Power Autom Equip, Nanjing. 2004;24(4):70–3, (in Chinese).
8. Ferkovic L, Ilic D, Malaric R. Mutual inductance of a precise rogowski coil in dependence of the position of primary conductor. IEEE Trans Instrum Meas, USA. 2009;58(1):122–8, (in Chinese).
9. Rezaei-Zare A, Ravani R, Sanaye-Pasand M, et al. An accurate CT model based on preisach theory for the analysis of electromagnetic transients, USA. IEEE Trans Power Deliv. 2008;23(1):233–42.
10. Sunan L, Chaobo T, Xicai Z. Performance analysis of air-core CT. Proc CSEE, Nanjing. 2004;24(3):108–13, (in Chinese).

Chapter 13
The Application of Fuzzy Pattern Recognition on Electromotor Malfunction Diagnosis

Tan Xia and Juan Wang

Abstract Until now, the fault detection on running motor is deemed a complicated and uncertain problem. In this chapter, a method for motor fault diagnosis based on fuzzy pattern recognition is proposed. With the fuzzy method, the mathematics model and the membership function of diagnosis are presented to provide a way for motor fault diagnosis.

Keywords Fuzzy logic · Pattern recognition · Electromotor malfunction · Diagnostic matrix

13.1 Introduction

When a fault happens, both financial losses and undesired situations may occur. In this sense, it is important to detect the incipient faults and diagnose which faults have occurred [1]. Until now, the fault detection on a running motor is deemed a complicated and uncertain problem. The qualitative decision of the motor failure is not accurate to solve problems timely, and classical mathematical methods are powerless here [2].

Many studies have focused on the analysis of the stator current. These methods are more accurate in diagnosis; however, the cost is high, and the type of fault diagnosis is limited.

The actual work primarily depends on the experience of engineers by means of observation, reasoning, and logic analysis such as excessive noise, vibration, strong or overheating signs to infer. The experience and the reasoning are full of

T. Xia (✉)
Wuhan Vocational College of Software and Engineering, 430205 Wuhan, China
e-mail: tansia@163.com

J. Wang
Wuhan No.2 Light Industry School, 430080 Wuhan, China
e-mail: wj810701@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_13

ambiguity. With the fuzzy method, the mathematics model and the membership function of diagnosis are presented to provide a way for the purpose of electromotor malfunction diagnosis.

13.2 Fuzzy Diagnostic Model

Definition 1 Assume the occurrence of certain failure, *m* kinds of fault symptoms occur with varying degrees, described as sign of the *i*th variable; thus, the fault symptoms fuzzy vector is proposed:

$$x_i(i=1,2,\cdots,m),=(x_1,x_2,\cdots,x_m),x_i \in [0\ ,1]. \tag{1}$$

When $x_i = 0$ *or* 1, expresses the sign—presence or absence.

Definition 2 Let the failure mode identified by *n* classes; thus, the variable model *j* is described as:

$$\begin{aligned} &y_j(j=1,2,\cdots,n),\underline{Y}=(S_{Y_1}(y_1),S_{Y_2}(y_2),\cdots,S_{Y_n}(y_{\mathrm{n}})) \\ &S_{Y_j}\in[0,1],\underline{Y}=(y_1,y_2,\cdots,y_{\mathrm{n}}). \end{aligned} \tag{2}$$

Definition 3 $S_{y_j}(x_i)=r_{ij},(i=1,2,\cdots m;\ j=1,2\cdots,n)$, as the fuzzy membership of the ith failure symptom x_i in the *j–th* failure mode. $\underline{R}=(r_{ij})_{m\times n}, r_{ij}\in[0,1]$ is the fault fuzzy diagnosis matrix:

$$\underline{R}=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix}. \tag{3}$$

13.3 Synthesis of Fuzzy Relations

Relationship based on fuzzy synthetic transformation, X and R synthesis operation:

$$\underline{Y}=\underline{\mathrm{X}}\circ\underline{R}=(x1,x2,\cdots,x_m)\circ\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \tag{4}$$

For $\underline{Y}$, the synthetic computing model can be used as below:

a. $M(\wedge,\vee), y_j = \underset{i=1}{\overset{m}{\vee}}(x_i \wedge r_{ij})$.

After obtaining the data of minimum and maximum, synthesizing operation of the model use the principle of maximum and minimum [3]. The main feature is the prominent role of the fault symptom factor without considering the impact of other signs of factors.

b. $M(\wedge,\oplus), y_j = \min\{1, \sum_{i=1}^{m} x_i \wedge r_{ij}\}$.

The model takes into account the minimum degree of membership of each fault symptom of failure modes, characterized by prominent signs of the impact of the secondary fault.

c. $M(",+), y_j = \sum_{i=1}^{m} x_i \bullet r_{ij}$.

This model is a weighted average of the model-based fault diagnosis in both the primary and the secondary factors at the same time with the effect of all factors of failure symptoms taken into consideration.

The chapter uses this model for fault pattern recognition in order to further improve the diagnostic accuracy; besides, the model can also be used in three different results mentioned above. Then it analyzes and compares the results.

13.4 Determine the Membership Function

In this chapter, we determine the fuzzy membership matrix of the fault diagnosis method by means of empirical methods and statistical combination.

The appearance of various types of the failure probabilities of each symptom is given by the theoretical analysis and empirical knowledge. We descript the degree of membership l_{ij}.

Definition 4: Each fault symptom corresponds to the meaning of failure modes membership: $l_{ij}=1.0$, must appear; $l_{ij}=0.8$, which is more likely to occur; $l_{ij}=0.6$, which is likely to occur; $l_{ij}=0.4$, which is somewhat likely to occur; $l_{ij}=0.2$, which is unlikely to occur; and $l_{ij}=0$, which is impossible to occur.

The membership is determined by the statistical data:

$$t_{ij} = \frac{\text{The number of presence of section } I \text{ in the } J-\text{th failure mode}}{\text{The total number of the presence in class } I}. \quad (5)$$

The empirical method weights w_1; the statistics act weights w_2;

$$0 \le w_1, w_2 \le 1, w_1 + w_2 = 1.$$

The fuzzy membership comprehensive experience and statistical methods are adopted to determine the fuzzy membership r_{ij}:

$$r_{ij} = w_1 l_{ij} + w_2 t_{ij}, i = 1, 2, \cdots, m; \; j = 1, 2, \cdots n. \tag{6}$$

Thus we can get the fuzzy diagnosis matrix $\underline{R} = (r_{ij})_{m \times n}$.

13.5 Principle of the Fuzzy Diagnosis

Sign the input vague signs vector $\underline{X} = (x_1, x_2, \cdots, x_m)$. Upon normalization, it has the right role model in the weighted average, $x_i' = x_i / \sum_{i=1}^{m} x_i$.

The failure modes can be obtained based on the relationship between the synthesis of the fuzzy vector $\underline{Y} = (y_1, y_2, \cdots, y_n), y_f = \max\{y_j \mid j = 1, 2, \cdots, n\}$. Inferred by the principle of maximum degree, y_f is the failure category (Fig. 13.1).

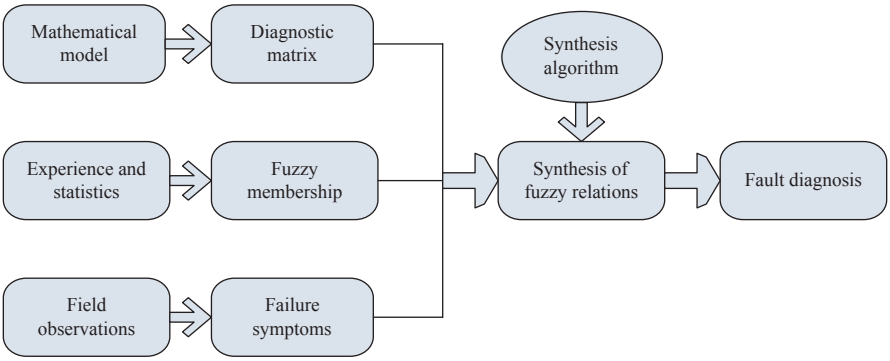


Fig. 13.1 Schematic motor fault diagnosis of the fuzzy model

13.6 Example Diagnosis

13.6.1 Selection of the Fault Symptoms and the Fault Mode

In the field of the induction motor diagnosis, the fault symptoms and the failure modes can be described as to the failure symptoms (operating condition monitoring) as given below:

x_1 (abnormal noise), x_2 (vibration increases), x_3 (local overheating), x_4 (three-phase current asymmetry), x_5 (speed down), and x_6 (corona).

As to the failure mode:

y_1 (turn short circuit between the stator), y_2 (loose stator core), y_3 (insulation aging), y_4 (cable damage), y_5 (end ring fracture), and y_6 (bearing failure).

13.6.2 Determination of the Fault Diagnosis Matrix

According to the theoretical analysis and the empirical knowledge, the fuzzy membership is given by definition 4 between the fault symptoms and the fault modes, which are shown in Table 13.1:

Previous experience of practical work and the existing statistics are collated in Table 13.2:

Table 13.1 can give the fuzzy relationship matrix of experience directly:

$$\underline{L} = \begin{bmatrix} 0.4 & 0.8 & 0.4 & 0.6 & 0.6 & 0.8 \\ 0.8 & 1.0 & 0.4 & 0.4 & 0.8 & 1.0 \\ 1.0 & 0.6 & 0.8 & 0.2 & 0.2 & 0.6 \\ 0.8 & 0.4 & 0.2 & 0.8 & 0.4 & 0.4 \\ 0.2 & 0.6 & 0.2 & 0.4 & 0.2 & 0.8 \\ 0.4 & 0.2 & 0.8 & 0.8 & 0.2 & 0.2 \end{bmatrix} \tag{7}$$

Table 13.1 Induction motor fuzzy relationship with the signs of failure modes

Fault symptoms	Fault modes y_j					
x_i	y_1	y_2	y_3	y_4	y_5	y_6
x_1	0.4	0.8	0.4	0.6	0.6	0.8
x_2	0.8	1.0	0.4	0.4	0.8	1.0
x_3	1.0	0.6	0.8	0.2	0.2	0.6
x_4	0.8	0.4	0.2	0.8	0.4	0.4
x_5	0.2	0.6	0.2	0.4	0.2	0.8
x_6	0.4	0.2	0.8	0.8	0.2	0.2

Table 13.2 Induction motor fault symptom and failure modes of membership statistics

Symptom *xi*	Fault *j*					
	y_1	y_2	y_3	y_4	y_5	y_6
x_1	5	3	1	3	4	6
x_2	3	4	0	1	2	6
x_3	6	3	5	0	2	1
x_4	6	2	0	7	1	0
x_5	2	1	0	2	0	2
x_6	1	0	3	3	0	1

Table 13.2 can give the fuzzy relationship matrix of statistics:

$$\underline{T} = \begin{matrix} 0.227 & 0.136 & 0.045 & 0.136 & 0.182 & 0.273 \\ 0.188 & 0.25 & 0 & 0.063 & 0.125 & 0.375 \\ 0.353 & 0.176 & 0.294 & 0 & 0.118 & 0.059 \\ 0.375 & 0.125 & 0 & 0.438 & 0.063 & 0 \\ 0.286 & 0.143 & 0 & 0.286 & 0 & 0.286 \\ 0.125 & 0 & 0.375 & 0.375 & 0 & 0.125 \end{matrix} \tag{8}$$

As the motor is a typical nonlinear system, it is difficult to obtain the diagnosis samples [4]. Several works have been focused on detecting early mechanical and electrical faults before damage appears in the motor. However, the main drawback of them is the complexity on the motor's signal mathematical processing [5]. Due to the small statistical sample, let the weight of experience $w_1 = 0.7$; statistical weights $w_2 = 0.3$;

then $r_{ij} = 0.7l_{ij} + 0.3t_{ij}$, the fuzzy diagnosis matrix is given as:

$$\underline{R} = \begin{bmatrix} 0.3481 & 0.6008 & 0.2935 & 0.4608 & 0.4746 & 0.6419 \\ 0.6164 & 0.7750 & 0.2800 & 0.2989 & 0.5975 & 0.8125 \\ 0.8059 & 0.4728 & 0.6482 & 0.1400 & 0.1754 & 0.4377 \\ 0.6725 & 0.3175 & 0.1400 & 0.6914 & 0.2989 & 0.2800 \\ 0.2258 & 0.4629 & 0.1400 & 0.3658 & 0.1400 & 0.6458 \\ 0.3175 & 0.1400 & 0.6725 & 0.6725 & 0.1400 & 0.1775 \end{bmatrix} \tag{9}$$

The site monitoring shows signs of motor failure, such as the presence x_1, x_3, and x_4. We enter the fault symptoms vector X = (1 0 1 1 0 0). After normalization, the result can be obtained according to the fuzzy relations:

$$\underline{Y} = (0.6082\ 0.4632\ 0.3602\ 0.4303\ 0.3160\ 0.4527) \tag{10}$$

Based on the principle of maximum degree of membership, we can judge the diagnostic fault mode as y_1, the stator inter-turn short circuit.

13.7 Conclusion

On the basis of theoretical analysis, the mathematical model and combined diagnostic examples can explain the fuzzy pattern recognition method of the motor fault diagnosis, which is feasible so as to provide a way for the computer-aided diagnosis of motor failure. This method can also be combined with the diagnostic expert system to obtain a more effective application of the high degree of complexity in large-scale dynamic systems.

This is a practical example of motor fault diagnosis, which is on the cold-rolled silicon steel production line of Wuhan Iron and Steel *Company (*WISCO) in 2013; however, this method is simulated in the motor laboratory of two schools according to the simulation of the broken rotor bars and eccentricity fault. Our method is proved to be effective. While compared with many kinds of online detection methods based on stator current, the fault detection accuracy rate is low. However, it can detect more types of faults with simple operation and low cost.

The study requires more follow-up samples, so we can determine more practical membership functions. This method can also be expanded to some kinds of fuzzy clustering analysis in order to diagnose the motor failure more accurately.

References

1. Ali U, Raif B. Real-time condition monitoring and fault diagnosis in switched reluctance motors with Kohonen neural network. J Zhejiang University SCI. C. December 2013;14(12):941–52.
2. Kyusung K, Alexander GP. Induction motor fault diagnosis based on neuropredictors and wavelet signal processing. IEEE Trans. 2002;7(2):201–19.
3. Wang XZ. Fuzzy spatial information processing. Wuhan: Wuhan University Press; 2003. pp. 1–266 (In Chinese).
4. Silva JL, Cardoso AJ. Bearing failures diagnosis in three-phase induction motors by extended Park's vector approach. The conference of 2005 industrial electronics society. IECON. 2005;28(4):648–53.
5. José JC, Luis PS, Victor ML, José JM. Misalignment identification in induction motors using orbital pattern analysis. Progress in pattern recognition, image analysis, computer vision, and applications lecture notes in computer science volume 8259; 2013; pp. 50–8.

Chapter 14
Low Speed Stability of Tracking Turntable for the Unmanned Air Vehicle Landing

Chengzhi Su, Yiping Ma, Yanjing Wang, Yuan Fang and Demin Wang

Abstract In the unmanned air vehicle (UAV) landing test, the rate fluctuation of tracking turntable appears in using the grounding active laser guidance system. To improve its stability, the reasons for tracking turntable rate fluctuation at low speed are analyzed. It is verified that the motor cogging torque fluctuation is the main reason at low speed by modeling and experiment. A disturbance observer based compensation control on the basis of the traditional double closed loop PID correction is proposed. Through analyzing the dynamic structure of tracking turntable, the compensation equation of disturbance torque described by the angular acceleration and torque current is derivated and then the torque disturbance observer structure is designed based on the equation. The comparative test shows that the rate fluctuation is 2.12 % after compensation when the speed is 1(°)/s, lower over 3 times than by only the traditional double closed loop PID correction. This method proposed can improve low speed stability of tracking Turntable for the UAV landing.

Keywords Unmanned air vehicle · Landing guidance · Tracking turntable · Rate fluctuation · Disturbance observer

Y. Ma (✉) · C. Su · Y. Wang · Y. Fang · D. Wang
College of Mechanical and Electrical, Changchun University of Science and Technology University, 130000 Changchun, China
e-mail: sunnypp3@163.com

C. Su
e-mail: chengzhi_su@126.com

Y. Wang
e-mail: 1040286587@qq.com

Y. Fang
e-mail: 305175945@qq.com

D. Wang
e-mail: ccwdm@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_14

14.1 Introduction

To improve the reliability of UAV landing on the ground[1], the grounding active laser guidance system has received widely theoretical and technical research and applicational development from home and abroad owing to its characters of strong anti-jamming capability, no relying on external factors, etc [2–3]. This landing guidance system adopts laser to range, two-dimensional turntable tracking UAV to measure angle and offers real-time UAV landing information [4–6]. In the process of leading the test, rate fluctuation of the turntable appears and it affects the tracking precision. Domestic and foreign scholars have done a great deal of work in the research of low-speed stability of turntable. Document [7] proposed a adaptive control strategy based on system identification to solve the problem that the motor torque fluctuation of low-speed servo system influences speed stability; Documents [8–10] proposed the repetitive control method to realize effective inhibition of such fluctuations; Documents [11–12] present the method of using neural network to inhibit fluctuant torque; Document [13] analyzed the effect of turntable`s rate fluctuation, and proposes a adaptive inhibition method to rate fluctuation of position sensor`s periodic error. Based on the test, the reasons for tracking turntable rate fluctuation at low speed are analyzed, and a disturbance observer-based compensation control on the basis of a traditional PID compensation control method is proposed, and it turns out that this method can effectively reduce the negative impact of the motor torque on rate stability.

14.2 Tracking System Model and Analysis of Rate Fluctuation

This two dimensional turntable adopts permanent-magnet DC torque motor, and its number of magnetic poles is 14, tooth number N is 85. First, mathematical model of the system is established to analyse the low-speed stability property of tracking turntable and then the dynamic structure diagram is shown in Fig. 14.2.

Motor has two input, one is the control input rectifier voltage Ud, and the other one is equivalent disturbance torque current $\mathrm{I_{dL}} = \mathrm{T_L}/\mathrm{C_m}$, T_L is disturbance torque, n is motor speed, E is armature voltage, $T_\mathrm{L} = \mathrm{L}/\mathrm{R}$ is electromagnetic time constant of armature circuit, $T_m = (4gJR)/(375C_eC_m)$ is tracking system mechanical electrical time constant. Table 14.1 is systematic parameters of tracking turntable (Fig. 14.1).

On account of the unequal distribution, reversing, the cogging effect and other problems of armature winding, DC torque motor inevitably produces torque fluctuation in the process of movement and the motor torque fluctuation has larger influence on rate stability at low speed. Because of low work rate and short journeys, the torque fluctuation mainly is high frequency fluctuations which are related to tooth number, and the mathematical model of the fluctuating torque $\mathrm{f_r}$ can approximately be sine function as follows:

$$\mathrm{f_r} = A\sin(2\pi \mathrm{ft} + \phi) \tag{14.1}$$

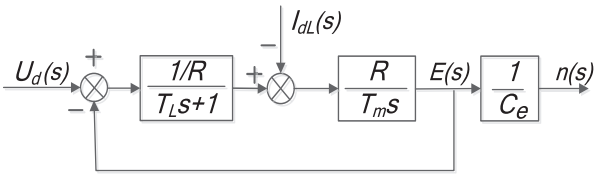


Fig. 14.1 Tracking system model

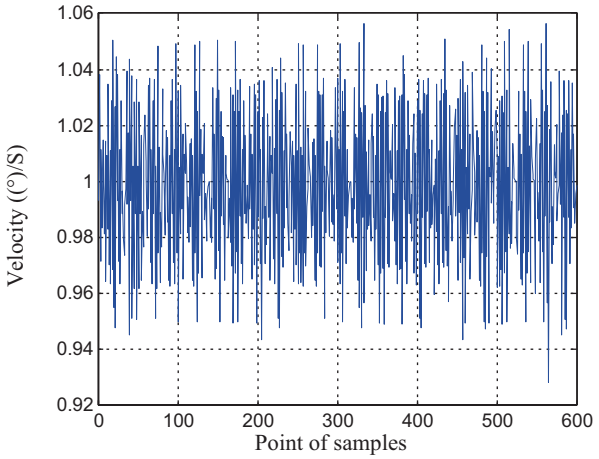


Fig. 14.2 Rate fluctuation at ω = 1(°)/s

Table 14.1 Tracking turntable system parameters

Symbol	Name of parameters	Value	Unit
L	Armature inductance	0.0159	mL
R	Armature resistance	9.3	Ω
J	Moment of inertia	950	$kg.m^2 \times 10^{-5}$
C_e	Back electromotive force coefficient	0.25	$v/r.min^{-1}$
C_m	The torque current ratio	2.39	Nm/A

A is the amplitude of harmonic fluctuation torque, $f = N/(360\omega)$ is the frequency of harmonic fluctuation torque, N is tooth number, ω is rotating angular rate of tracking system, φ is the phase angle related to the initial position.

Figure. 14.2 is the rate curve when ω is 1(°)/s, rate fluctuation is 7.22%. The spectrum, shown in Fig. 14.3, is obtained by removing the DC component and then making Fourier transform to the data in Fig. 14.2. According to Fig. 14.3, we know that: the rate of tracking system is approximately in the form of periodic fluctuation, the frequency of rate fluctuation is mainly at 0.233 Hz. The theoretical value of the cogging torque fluctuation is 0.236 Hz, indicating that the actual rate fluctuation is consistent with the theory. Change the speed to 2(°)/s, 3(°)/s, the frequencies of rate fluctuation are 0.469 Hz, 0.703 Hz, approximately satisfied Eq. (14.1). All of those show that unequal distribution of the cogging torque is the main reason for the rate fluctuation of turntable and Eq. (14.1) is right.

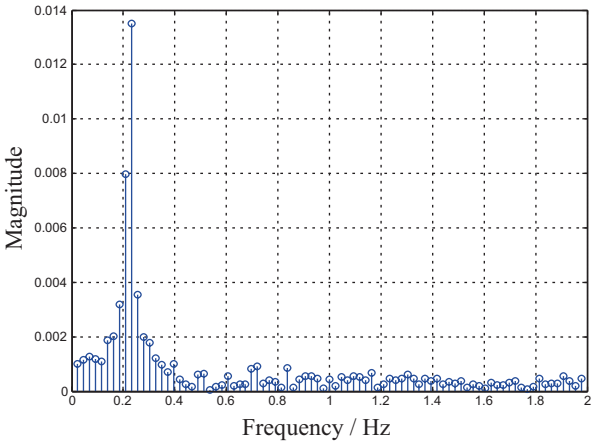


Fig. 14.3 Spectrum of rate at ω = 1(°)/s

14.3 Traditional Double Closed Loop PID Correction

This correction system adopts the current ACR, speed ASR, the double closed loop controlling structure with two independent regulators, as shown in Fig. 14.4. $K_s/(T_s+1)$ is the equivalent transfer of tracking system based on PWM actuator, $T_s = 0.004$ is the time constant of PWM controller, $K_s = 2.4$ is the gain of PWM controller, β is feedback coefficient of current loop, α is feedback coefficient of speed loop.

To make the current loop have good following performance, I-type system is used to correct this loop, ACR regulator adopts PI, and its transfer function is:

$$W_{ACS} = (K_I + K_P s)/s \tag{14.2}$$

K_P scale coefficient of current regulator

K_I integral coefficient of current regulator

To make the speed loop have no static error, II-type system is used to correct this link, ASR regulator adopts PI structure, and its transfer function is:

$$W_{ASR} = K_n(\tau_n s + 1)/(\tau_n s) \tag{14.3}$$

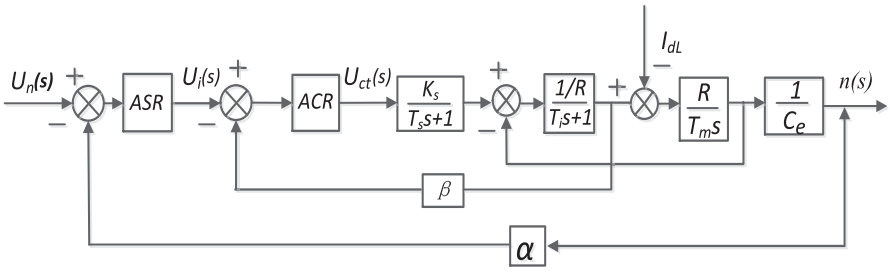


Fig. 14.4 traditional doubleclosed loop PID correction dynamic structure

K_n scale coefficient of speed regulator
τ_n lead time constant of speed regulator

In the circumstance of matlab/Simulink, establishing the simulation model of double closed loop PID controlling system, according to the harmonic peaks M_{rmin} minimum standards from modern control theory "oscillation index method" [14–15], the parameters are optimized as followed: current regulator ASR: $K_P = 1.6$, $K_I = 964.5$; speed regulator ACR: $K_n = 23.06$, $\tau_n = 0.06$; feedback coefficient of the current loop $\beta = 1$, feedback coefficient of the speed loop $\alpha = 1$. When the given rate is 1(°)/s, rate step response curve without the torque disturbance, shown in Fig. 14.5, is obtained by traditional double closed loop PID correction.

Figure 14.5 indicates that the tracking system has a better rapidity and stability with traditional double closed loop PID correction.

The solid line with torque disturbance in Fig. 14.6 is the rate step response curve adapted by traditional double closed loop PID correction. As known in Fig. 14.6, the simulation rate fluctuation is 1.9%. From Fig. 14.2, the real rate fluctuation is 7.22%. According to the requirement of UAV guidance, when the lowest tracking rate is 1(°)/s, rate fluctuation is less than 3%. The rate fluctuation requirement is not satisfied only by traditional double closed loop PID correction. It is clear that the traditional double closed loop PID correction is lack of effective inhibition to torque fluctuation.

14.4 Disturbance Observer Based Compensation Control

In order that the tracking system has not only better rapidity and stability, but also has robustness to torque fluctuation, according to modern control theory, we add the disturbance observer to compensate the effect of torque fluctuation on the basis of traditional double closed loop PID correction.

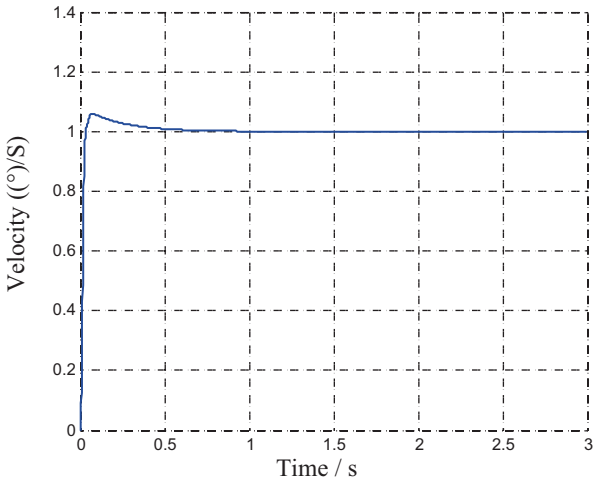


Fig. 14.5 Rate step response adapted by traditional double closed loop PID correction without the torque disturbance

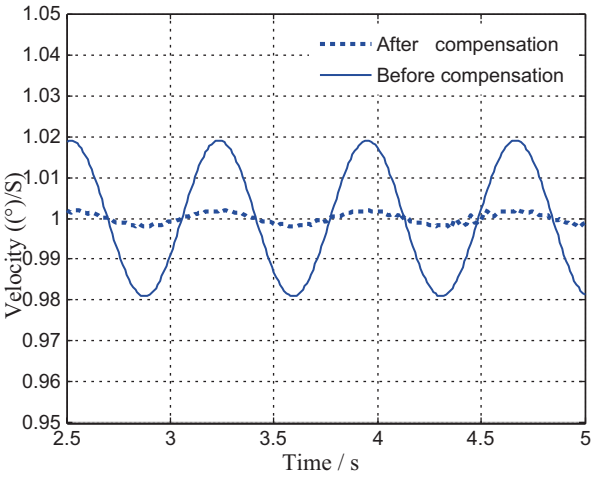


Fig. 14.6 Rate step response adapted by traditional double closed loop PID correction with the torque disturbance

According to the object and disturbance properties, we can get the new dynamic structure of tracking system just as the Fig. 14.7 and we know that:

$$I_d(s) - I_{dL}(s) = \frac{T_m C_e}{R}\dot{n}(s) \tag{14.4}$$

$$I_{dL}(s) = I_d(s) - \frac{T_m C_e}{R}\dot{n}(s) \tag{14.5}$$

$$I_{dL}(s) = I_d(s) - a{\cdot}s{\cdot}\dot{n}(s) \tag{14.6}$$

$$a = \frac{T_m C_e}{R}$$

The disturbing torque can be estimated by torque current and angular acceleration according to Eq. (14.6). We can get the angular acceleration by differential computing the angular rate, and adopt the low-pass filter to eliminate the differential noise of angular rate signal. We get the estimated compensation equation of disturbing torque I^*_{dL}:

$$I^*_{dL}(s) = \frac{1}{Ts+1}[I_d(s) - a{\cdot}s{\cdot}\dot{n}(s)] \tag{14.7}$$

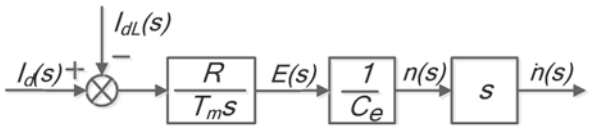


Fig. 14.7 Tracking system disturbance structure

Add the disturbing torque compensation structure in Fig. 14.8 into Fig. 14.4. We adopt the same simulated parameters which are in the above the traditional double closed loop PID correction in the circumstance of Simulink. The dotted line in Fig. 14.6 is the rate fluctuation curve after compensation, and the rate fluctuation is0.2 %, decreased nearly 10 times than before compensation.

14.5 Test and Analysis

The ground tests are done with the grounding active laser guiding system. When testing, the pitching axis of tracking turntable points to 30(°) relative to the zenith, azimuth axis rotates continuously at 1(°)/s, and the spectrum of disturbance torque is shown in Fig. 14.3. Collect real-time rate data, sample period is 0.01s, sample length is 72,000, and exclude the partial data at the beginning and the end to avoid the effect of acceleration and deceleration, retaining the effective sample length 60,000, and part of the data image is shown in Fig. 14.9, part of the data is shown in Table 14.2. Comparing the result in Table 14.2, we can know that when the V_{ave} of average rate is 0.9994(°)/s, Vp-p of rate fluctuation is 2.12 % after plus disturbance observer based compensation control, less 3 times than only by the traditional double closed loop PID correction.

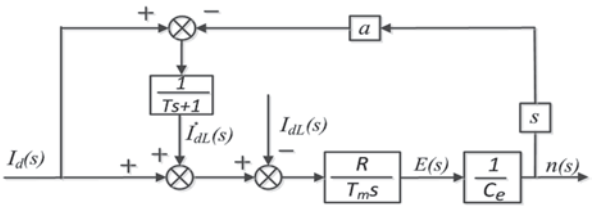


Fig. 14.8 Torque disturbance compensation structure

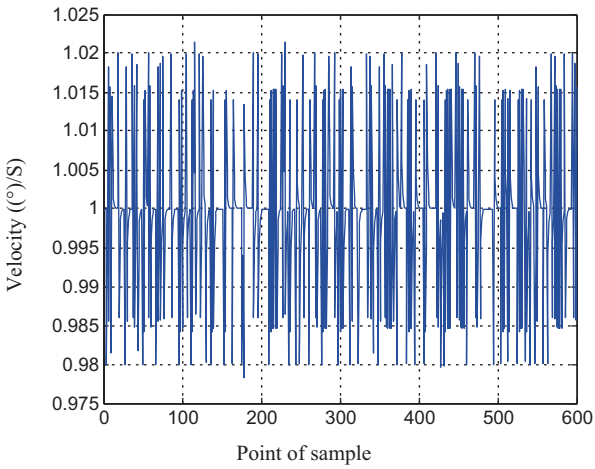


Fig. 14.9 Real rate curve after compensation

Table 14.2 Part of rate sampling data

No.	Only by the traditional double closed loop PID correction (°)/s	After plus disturbance observer based compensation control (°)/s	No.	Only by the traditional double closed loop PID correction (°)/s	After plus disturbance observer based compensation control (°)/s
1	0.993316892	0.999848209	17	1.029002177	1.000010514
2	1.037995068	0.999954464	18	0.968700653	1.000003154
3	0.971398524	0.999986339	19	1.050610196	1.020000946
4	1.011419556	0.979995902	20	1.026554918	0.995800085
5	0.983425867	0.993998773	21	1.044389943	0.999622008
6	0.993316892	0.999848209	22	0.993316983	0.999886602
7	1.004508328	0.985459889	23	1.037995095	0.999965981
8	0.981352498	1.015637967	24	0.971398528	0.999989794
9	1.014405756	1.004691393	25	1.011419559	0.979996938
10	0.964321725	0.981407417	26	0.963425868	0.993999081
11	1.009296517	1.014422225	27	1.02902776	1.018199724
12	0.982788955	1.004326668	28	0.988708328	0.985459917
13	1.034836687	1.001298445	29	0.996612498	0.995637975
14	0.970451006	1.00038943	30	0.978983753	0.998691393
15	1.011135302	1.000116822	V_{ave}	0.998528305	0.999484295
16	0.963340591	1.000035046	V_{p-p}	7.19%	2.12%

14.6 Conclusion

After analysing the reasons for rate fluctuation of tracking turntable at low speed, a disturbance observer based compensation control on the basis of the traditional double closed loop PID correction is proposed. And the test shows that:

1. In the UAV landing test, the motor cogging torque fluctuation is the main reason for causing rate fluctuation of tracking turntable at low speed which is used on the grounding active laser landing system;
2. When turntable is at the speed of 1(°)/s, the rate fluctuation is 2.12 % after adopting the disturbance observer based compensation control, lower over 3 times than only the traditional double closed loop PID correction.
3. The disturbance observer based compensation control proposed can improve the rate stability of tracking turntable.

References

1 Yang L. UAV landing guidance technology. Radio Eng China. 2009;39(12):29–31 (in chinese).
2 Can Diao, Yingxun Wang Yingxun, Wang J et al. Computer vision assisted autonomous landing of UAV. Acta Aeronoutica et Astronautica Sinica. 2008;29(2)supplement:79–85 (in chinese).
3 Xue D. Optical equipment subsystem tentative project research for passive UAV's optical auto-landing system. Ome Inf. 2007;28(2):6–12 (in chinese).
4 Cai Y, Wang G, Li G. An emergency UAV landing method based on ratio and laser:China, CN 103149938 A[P]. 2013–06–12
5 Hongyi Bu. Application of lidar in autonomous planetary landing. INFRARED. 2008;29(7):356–8 (in chinese).
6 Xu G, Cheng Y, Shen C. Unmanned air vehicle's navigation and automatic accurate landing in all weather based on infrared laser scan and computer vision. Acta Aeronoutica et Astronautica Sinica. 2004;25(5):499–504 (in chinese).
7 Song Y, Gao H, Zhang S et al. Adaptive compensation of torque ripple in DC torque motor. Opt Precis Eng. 2010;20(10): 18–9 (in chinese).
8 Liu J, Er L. Design of high precision digital repetitive controller for flight simulator servo system [J]. Acta Aeronoutica et Astronautica Sinica. 2004;25(1):59–61 (in chinese).
9 Liu Y, Zeng M, Su B. Robust repetitive control for improving rate smoothness of test turntable. Trans Nanjing Univ Aeronaut & Astronaut. 2005;22(2):19–21 (in chinese).
10 Huo X, Yao Y. Suppression of moment fluctuation in high precision rate table servo system. Electr Mach Control. 2007;11(2):27–9 (in chinese).
11 Zheng D, Na J, Ren X et al. Adaptive control of robotic servo system with friction compensation. IEEE Conference on robotics, automation and mechatronics (RAM), IEEE. 2011. p. 285–290.
12 Yi G, Ma G, Wang C. The compensation of speed ripple caused by the angle-measuring error. Proc CSEE. 2002;22(11): 34–6 (in chinese).
13 Xu G, Liu Y, Zhang Y. Analysis of causes of velocity ripple of rotary table and discussion of an adaptive elimination method. Aviat Precis Manuf Technol. 2004;8(1):31–35 (in chinese).
14 Zhou D. Problems of oscillation index method in concepts and design. Inf Control. 1983;6(3):53–6 (in chinese).
15 Fang M. Oscillation index method of designing automatic regulating system. Inf Control. 1978;7(2):48–50 (in chinese).

Part II
New Energy and Smart Grid

Chapter 15
Development of a Multiple Stage Single-three-Phase Power Converter

Min Lei, Quanzhu Zhang, Mingyan Xu and Xiaolei Sun

Abstract The project studied and designed a set of multiple stage single-three-phase power converter by using modern power electronic technology and automatic control technology. The paper firstly introduced the design of the main circuit and control system and explained the working principle of them in detail and then introduced the design of driving circuit. The test results showed that the power converter realized the transformation from the single-phase voltage to three-phase voltage and the continuous adjustment of output voltage and frequency, which proved the correctness and feasibility of the design. The single-three-phase power converter has functions of transformer and converter. It has developmental prospects and considerable market value, and it can be used as a platform for experiment teaching and research in the field of power electronics technology and driving and controlling of motor.

Keywords Multiple stage power converter · Insulated gate bipolar transistor (IGBT) · Inverter circuit · Regulation of voltage and frequency · Drive circuit

15.1 Introduction

The main task of power electronics is to realize the conversion of electrical energy, and power conversion technology [1–2] is an important meaning to accomplish this task; therefore, it is the most basic and the most important core technology of power electronics. With the constant update and development of power electronics and control theory [3], and the improvement of microelectronics technology and the emergence of new power electronic devices, especially the rapid development of IGBT or metal–oxide–semiconductor field-effect transistor (MOSFET), the applications of power conversion technology expand unceasingly to increase the technology of regulation of frequency and voltage to a very high level and have advantages of high efficiency and high power factor. Power conversion technology

Q. Zhang (✉) · M. Lei · M. Xu · X. Sun
Institute of Information and Control Technology, North China Institute of Science and Technology, 065201 Beijing, China
e-mail: zhangquanzhu1965@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_15

is mainly for the effective conversion of voltage and frequency. In order to meet the requirements of different voltage parameters in practical applications, it is often necessary to transform a voltage format to another voltage format. This program is researched and designed based on the development ideas. The system consists of former power converter (AC/DC/AC1), backward power converter (AC2/DC/AC), and high-frequency transformer. System parameters are as follows: Input voltage is single-phase voltage of AC220 V/50 Hz, output voltage is continuously adjustable voltage of 0–380 V, frequency range is continuously adjustable of 0–60 Hz with 0.1 Hz for frequency regulation accuracy, and the output power is 3 KW. The single-three-phase power converter can convert constant single-phase input voltage of AC220 V into three-phase AC voltage with rated value of AC380 V with continuous adjustment of frequency and voltage, and it adopts high-frequency transformer to realize electrical isolation of input and output and make the whole system small size and lightweight for easy transportation and maintenance.

15.2 Main Circuit and Working Principle

The main circuit block diagram of the multiple stage single-three-phase power converter is shown in Fig. 15.1. The working principle is that the sine voltage of AC220 V/50 Hz gets DC of 300 V through single-phase rectifier and filtration of diode, which realizes high-frequency square wave voltage conversion from DC300 V to AC300 V/20 kHz through single-phase inverter of IGBT and thus realizes square wave voltage conversion from AC300 V/20 kHz to AC540 V/20 kHz through high-frequency inverter. Moreover, the outputs continuously adjust three-phase sine wave voltage through rectifier and filtration of fast diode, three-phase inverter of IGBT, and L–C three-phase sinusoidal filter. Continuous adjustment of the output voltage is dependent on IGBT single-phase inverter of the former power converter by adopting pulse width modulation (PWM) control method, and the control operation of variable voltage and variable frequency (VVVF) for motor load is dependent on IGBT three-phase inverter of the backward power converter. High-frequency transformer not only achieves the voltage step-up transformation, but also achieves the electrical isolation between former stage and backward stage. According to the operation principle of transformer, the volume of the transformer iron core is inversely proportional to the frequency of the input AC voltage. Since the

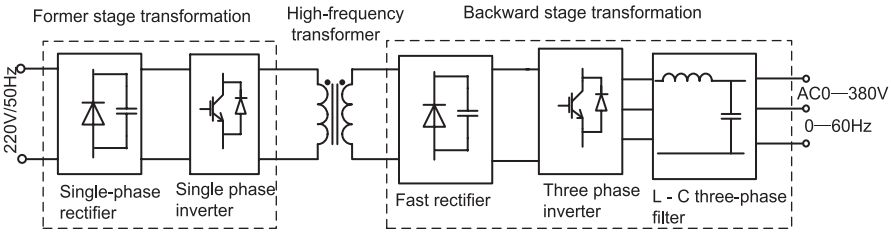


Fig. 15.1 Main circuit block diagram

input voltage of the transformer is 20 kHz high-frequency AC square wave voltage, it has adopted superfine crystal core to greatly reduce the size and weight of the converter. Former stage inverter is a single-phase inverter, and it selects a two-unit IGBT module of FF100R12KS4 as a master device with maximum withstand voltage of 1200 V, maximum current of 100 A, and switching frequency of 15–30 kHz. Three-phase inverter of backward stage inverter selects a two-unit IGBT module of BSM50GB120DN2 as a master device with maximum withstand voltage of 1200 V, maximum current of 50 A, and switching frequency of 10–20 kHz. Main circuit adopted modular design with stable circuit structure, realized the electrical isolation between input signal and output signal, and completed all the output requirements.

15.3 Control System Design

Control system of the multiple stage single-three-phase power converter makes SPMC75F2413 A from Sunplus75 Micro Control Unit (MCU) as the core coupled with a series of external circuits to achieve the effective control of output voltage and frequency. The principle block diagram is shown in Fig. 15.2. System detection module collects output three-phase voltage and current value and then sends to SPMC75F2413 A [4–5]; after processing, SPMC75F2413 A outputs the corresponding control signal to the integrated PWM control chip SG1525 for analysis and calculation to change the duty ratio of output pulse signal through the driver module and realizes the effective regulation of the output voltage. At the same time, the system detection module collects information such as voltage, current, and temperature of each link to realize condition monitoring, fault alarm, and protection function. SPMC75F2413 A is a kind of industrial 16-bit single chip with rich hardware resources, 80 I/O pins, watchdog, serial communication, 10 AD conversion circuit, etc. The single-chip microcomputer is used to implement operation and management of control system, data acquisition and display, closed-loop control, etc.

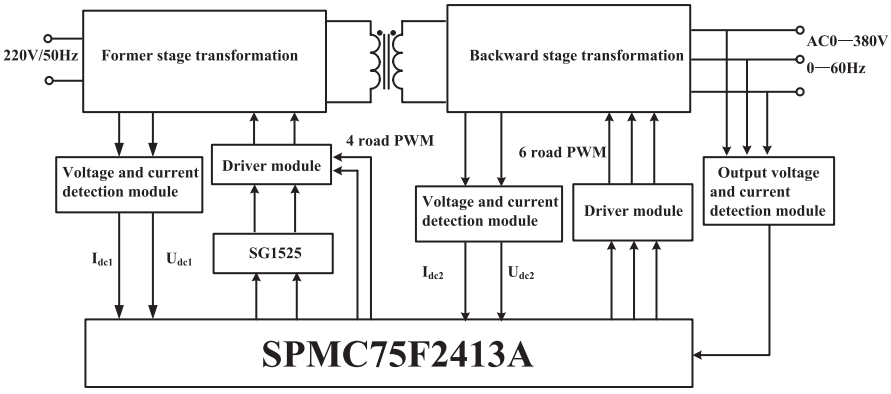


Fig. 15.2 Control system principle block diagram

15.3.1 The Design of Former Stage Driver Circuit

Former stage driver circuit [6–7] makes integrated ICPC929 and PC923 from SHARP Company as the core, and the principle diagram is shown in Fig. 15.3. The working principle is that the pulse signal produced by MCU is sent to the three feet of PC1; if the input signal is of low electrical level, the six feet of PC1 outputs high-level signal, and it provides positive bias current for Q2 through resistance R6. The conduction of Q2 makes positive power supply voltage that is sent to gate G1 of IGBT through the grid resistors R7 and R8. At the point, the IGBT is turned on. If the input signal of PC1 is of high electrical level, the six feet of PC1 outputs negative voltage, thus providing positive bias current for Q3 through resistance R6. The conduction of Q3 makes negative power supply voltage which is sent to gate G1 of IGBT through the grid resistors R7 and R8. At the point, the IGBT is turned off. C1 and C2 are used to absorb the voltage of the high-frequency harmonic. The 9 feet of PC2 is fault detection signal input pin for IGBT module, and it makes up IGBT conducting tube pressure drop detection circuit with external resistor R14 and diodes D1 and D2. High-speed optical coupling device P1 implements the electric isolation between the control signal and driving signal and implements the overcurrent protection of IGBT at the same time. The chips of PC929 and PC923 are suitable for small capacity of IGBT pulse drive, which are usually used in pairs. PC923 is generally chosen to drive the upper bridge arm, and PC929 is for lower bridge arm in the design.

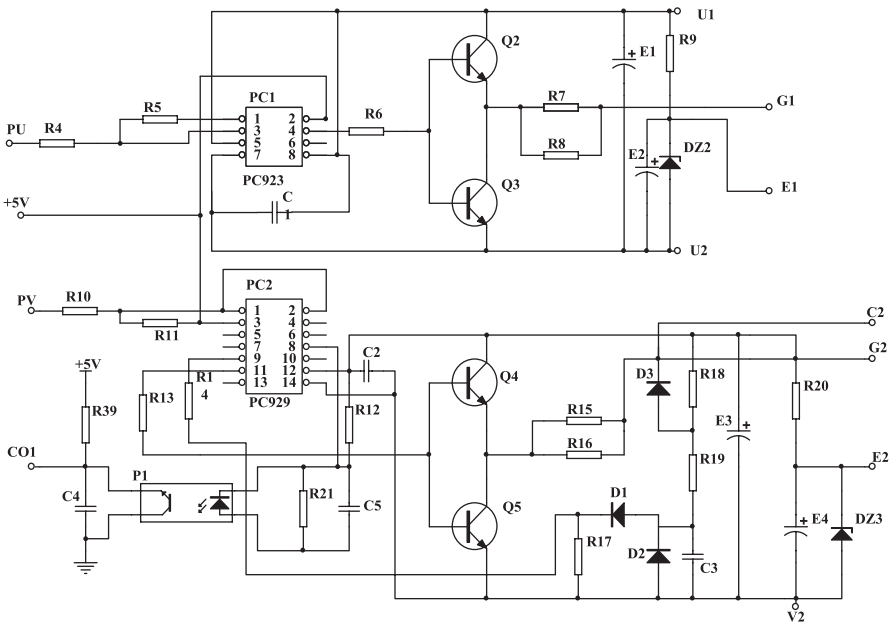


Fig. 15.3 The principle diagram of former stage driver circuit

15.3.2 The Design of Backward Stage Driver Circuit

Backward stage driver circuit [8] makes 2SD315 module from Swiss Concept Company as the core; its principle diagram is shown in Fig. 15.4. The driver circuit uses the operation of single power supply and positive and negative voltages to drive. PWM1 and PWM2 correspond, respectively, to control pulse input signal of two-unit IGBT from the same bridge and they are independent of each other with the dead time between them. SO1 and SO2 output two-state feedback signals and they are connected together to achieve protective function. The fault signal is outputted as long as the circuit is abnormal. SO outputs high when the circuit works normally. When a channel is detected to generate fault signal, SO outputs low. While the low level is sent to latch circuit, the latched signal blockades the output signal of PWM to achieve the purpose of protecting IGBT in real time. Setting reasonably the reference resistor R_{th} (R_{th1} and R_{th2}) and external capacitance of C (C1 and C2) pin is the prerequisite for correctly using the driver module. In order to minimize the loss for open up and shut off the IGBT, therefore, we respectively provide grid resistance (RG11, RG12, RG21, RG22). Diodes (D1 and D2) are used to prevent backflow of current from the collector of IGBT to the C-terminus of 2SD315. 2SD315 constituted mainly of laser direct imaging (LDI), pulse transformer device, IGD, and DC/DC conversion device; it has a compact shape size, simple external electric interface, high output current, high power, and good electrical isolation, and also it can drive directly IGBT module without pulsing a push–pull power amplifier circuit. In practical applications, the drive circuit is tried to avoid IGBT module and shorten the gate length between PWM pulse signal and the grid to increase anti-jamming performance of drive circuit.

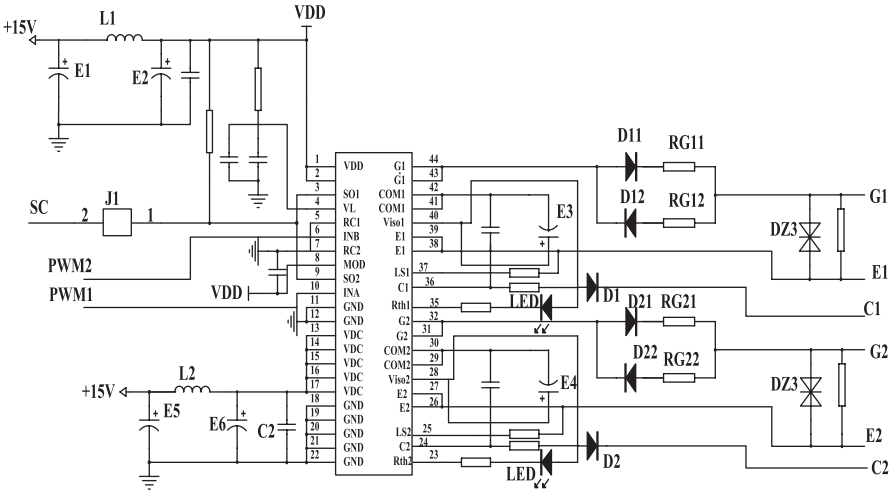


Fig. 15.4 The principle diagram of backward stage driver circuit

15.4 Results and Discussion

After testing step by step, we test the output function of the power inverter, mainly for its modulation function of frequency and voltage. The following shows a part of the voltage waveform by using oscilloscope under the different frequencies.

Setting the output frequency for 50 Hz and 60 Hz, the voltage waveform of U_{UV} before and after for three-phase LC filter is shown in Fig. 15.5a and b.

The waveform of two phase line voltages of U_{UV} and U_{VW} (after filtering) outputted from the power converter under rated voltage, and the frequency is shown in Fig. 15.6. The figure shows that the mutual difference between U_{UV} and U_{VW} is 120° with symmetry and balance between them, and it can drive three-phase asynchronous motor to run smoothly.

Figure 15.5a and b shows that the modulation function of frequency and voltage for the power converter is normal and the design requirements are met. Figure 6 shows that the output three-phase electrical phase angle difference of the power converter is 120° and very balanced.

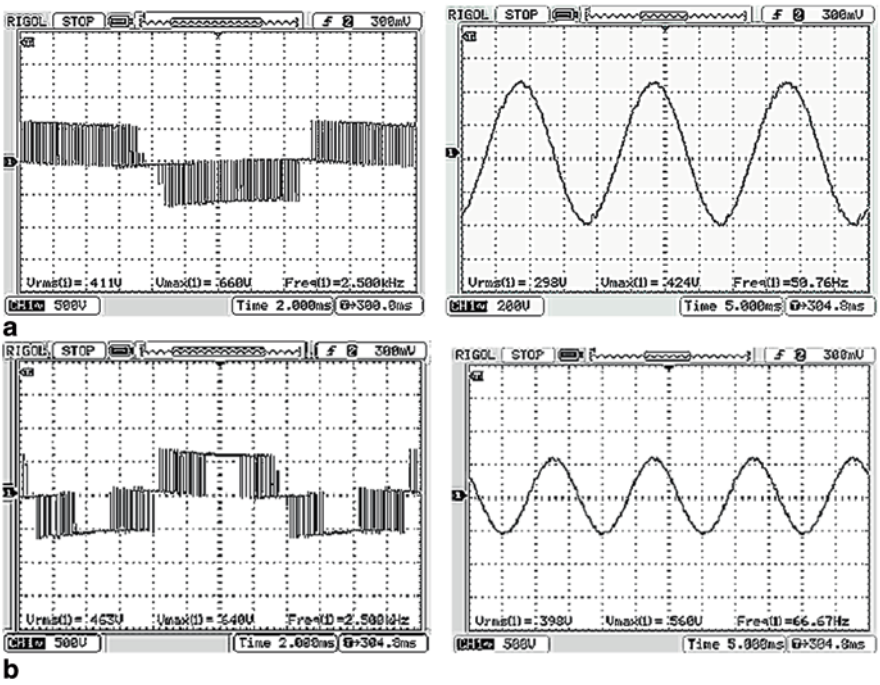


Fig. 15.5 **a** Voltage outputted from the inverter before and after the filter (f = 50 Hz). **b** Voltage outputted from the inverter before and after the filter (f = 60 Hz)

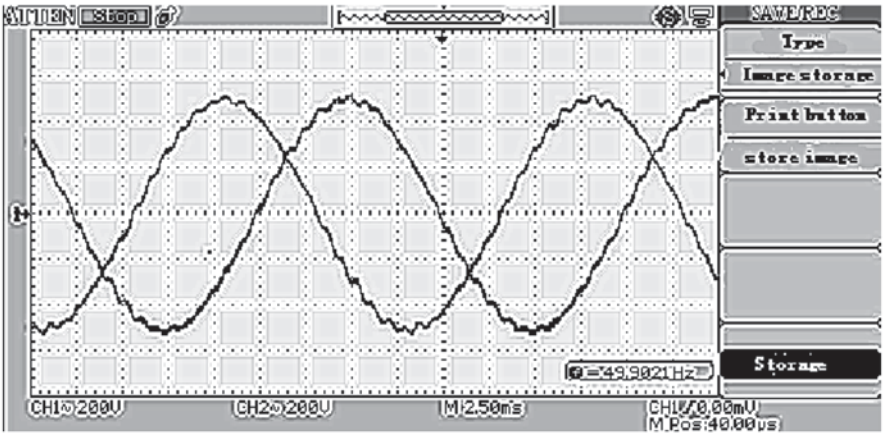


Fig. 15.6 The waveform of two phase line voltages of UUV and UVW

15.5 Conclusion

The main circuit of the multiple stage single-three-phase power converter studied and designed in this paper is made up of power frequency rectifier made of ordinary diode, former stage single-phase inverter, and backward stage three-phase inverter made of IGBT, which is a high frequency and fast rectifier made of fast recovery diode and high-frequency transformer of nanometer crystal core. The control system of the converter made SPMC75F2413A as the core combined with integrated PWM control chip SG1525. Integrated IC of PC929, PC923, and 2SD315 module makes up of former stage single-phase inverter driver circuit and backward stage three-phase inverter driver circuit. It proves through experiment the correctness and the feasibility of the system designs and achieves the desired design requirements. The power conversion system served as a typical application of electronic power conversion technology, it can be further developed for the application equipment and also act as a platform for experiment teaching and research in the field of power electronics technology for motor driving and controlling, and it can be popularized and applied in the industry.

References

1. Zhang X, Zhang J. Power converter application technology and practices. Beijing: China Electric Power Press; 2009. p. 135–41. (In Chinese).
2. Wang Z, Huang J. Power electronics. Beijing: China Machine Press; 2008. p. 97–106. (In Chinese).
3. Chen B, Chen M. AC adjusted speed system. Beijing: China Machine Press; 2005. p. 159–64. (In Chinese)

4. Chen L, Lv T, Wang K. Design of double closed loop DC speed control system based on SPMC75F2413. Comput Eng Softw. 2011;32(3):95–7. (In Chinese).
5. Li D, Huang C, Tang N, Chen G. Design and realization of DSP-based single-phase inverter with dual-closed loops control. Electrotech Electr. 2011;34(3):21–3. (In Chinese).
6. Hao Z, Wang H. Research of the phase shift FB-PWM converter based on DSP. Electr Drive. 2007;37(1):27–30. (In Chinese).
7. Zhu Y, Wang Q, Zhang Z, Li W. Simulation and experimental of reversible PWM rectifier based on double closed-loop control system. Coal Mine Mach. 2010;31(5):57–60. (In Chinese).
8. Li N, Zhang Q. Design of three-phase SPWM waveform generator based on SPMC75F2413A single-chip microcomputer. Electr Drive Autom. 2008;30(2):37–40. (In Chinese).

Chapter 16
Wind Farm Reactive Power Compensation Capacity Configuration

Chaoying Yang, Jinhao Wang, Huijie Du, Si Liu and Hao Huang

Abstract Reactive power compensation is an important technology to improve power quality of wind power. This paper analyzes various influence factors and reasonable arrangement for wind farm reactive power compensation capacity. Analysis and calculation are carried out for the inconsistent requirements in design regulations and influence factors for transmission line losses. At the same time, the configuration scheme for total compensation capacity, dynamic compensation device capacity, capacitor branches capacities, and their grouping modes are also given. The feasibility of compensation capacity configuration scheme is verified through multiple wind farm calculation cases. The purpose of flexible control of voltage and reactive power can be obtained with the reasonable capacity configuration for wind farm reactive power compensation devices.

Keywords Wind power generation · Reactive power compensation device · Dynamic compensation capacity · Power capacitor bank

16.1 Introduction

Recently, more and more attentions are paid to the study of wind farm reactive power compensation capacity configuration. Runqing Bai summarized the common compensation form of wind farm [1]. In consideration of the particularity of wind power generation, Yang Yu studied dynamic compensation devices to maintain wind farm stability [2]. Wei Gu compared characteristics of SVC and STATCOM used in wind farm, improvement of the steady and transient performance are verified [3].

C. Yang (✉)
State Grid Shanxi Electric Power Company, 030001 Taiyuan, China
e-mail: 1090474242@qq.com

J. Wang · H. Du
State Grid Shanxi Electric Power Research Institute, 030001 Taiyuan, China

S. Liu · H. Huang
North China Electric Power University, 102206 Beijing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_16

For reactive power compensation capacity design and configuration, fundamental regulation and scope had been described [4–7]. Qiang Zhang theoretically studied the reactive power capacity selection for wind farm [8]. However, the dynamic compensation capacity, capacitor branches capacities and their grouping modes had not been discussed in detail in the above references. As for different factors affecting the compensation capacity, Kai Wang made a detailed analysis [9]. But the capacity of dynamic compensation device is only sharing total compensation capacity. Yongwu Zhang established simulation models according to different wind turbines and output [10], but the selection and allocation of capacity is mainly determined by the simulation waveforms. However, there has not been a detailed method of dynamic and capacitor reactive power compensation capacity configuration about wind farm until now.

This paper discusses various influence factors and reasonable arrangement for wind farm reactive compensation capacity. Considering installed capacity and the length of transmission line, etc., a detailed calculation method of the total compensation capacity is given. Then, the selection criteria of dynamic compensation device capacity, capacitor branches capacities, and their grouping modes are proposed. The feasibility of the capacity configuration scheme is verified by wind farm calculation cases.

16.2 The Total Compensation Capacity for Wind Farm

There are many standards that involved wind farm reactive power compensation problems, and Q/GDW392-2009 and GB/T19963-2011 qualitatively provide the contents which should be included in the compensation capacity [4] [5]. Q/GDW212-2008 and Q/GDW392-2009 quantitatively regulate the reactive power compensation capacity [6] [7]. However, for a specific wind farm configuration, and contradictory or not according to these standards, still needs detailed calculation and analysis.

16.2.1 Reactive Power Loss of Main Transformer

Reactive power loss of main transformer is mainly divided into the no-load loss (constant loss) and load loss (variable loss), which can be calculated by the formula (1).

$$\Delta Q_{\mathrm{T}} = \Delta Q_0 + \Delta Q_{\mathrm{S}}\left(\frac{S^2}{{S_{\mathrm{N}}}^2}\right) = \frac{I_0\%}{100} S_{\mathrm{N}} + \frac{U_{\mathrm{S}}\%}{100}\left(\frac{S^2}{{S_{\mathrm{N}}}^2}\right) \tag{16.1}$$

where ΔQ_{T} is the total transformer reactive power losses (kvar), ΔQ_0 is the no-load loss (kvar), ΔQ_{S} is the load loss (kvar), S_{N} is the rated capacity (kVA), S is the ap-

parent power (kVA), *U*s% is the percentage of short-circuit impedance, and $I_{0\%}$ is the percentage of no-load current.

16.2.2 Reactive Power Loss of Box-Type Transformer

The reactive power loss of box-type transformer is also divided into two parts, which can be calculated by the Eq. (16.1). As the principle of one wind turbine one box-type transformer, every 50 MW installed capacity generally required 33 box-type transformers accessed.

16.2.3 Assemble Line Reactive Power Loss

The reactive power losses of assemble lines consist of the consumed reactive power of line equivalent reactance and the issued reactive power of the equivalent susceptance to ground (charging power), as it is shown in the formula (2).

$$\Delta Q = \Delta Q_{\mathrm{L}} + \Delta Q_{\mathrm{B}} = \left(P_1^2 + Q_1^2\right)\frac{X}{U_1^2} - \left(U_1^2 + U_2^2\right)\frac{B}{2} = 3I^2X - \left(U_1^2 + U_2^2\right)\frac{B}{2} \quad (16.2)$$

where ΔQ is the assemble line losses (kvar), ΔQ_{L} is the inductive reactive power loss (kvar), ΔQ_{B} is the charging power (kvar), X is the reactance (Ω), B is the equivalent susceptance (S), P_1 is the line active power (kvar), Q_1 is the line reactive power (kvar), I is the line current (A), and U_1 and U_2 are the beginning and terminal voltages of the line (kV).

16.2.4 Transmission Line Reactive Power Loss

Transmission line reactive power loss is also divided into two parts, which can be calculated by the formula (2). Among them, capacitive compensation capacity includes inductive reactive power loss, and dynamic compensation capacity includes charging power loss.

To analyze the problem, this paper takes typical wind farm cases as examples, which include 1.5 MW doubly fed induction wind turbines, 1.6MVA box-type transformers, 120MVA main transformer, LGJ-240 assemble lines, and LGJ-400 transmission line access to the grid. Tables 16.1 and 16.2 show the results under different conditions.

The total reactive power loss can be basically determined by the above four factors, as it is shown in Fig. 16.1.

Table 16.1 Installed capacity and reactive power loss of different parts

Installed capacity/MW	Main transformer loss/Mvar	Box-type transformer loss/Mvar	Assemble line loss/Mvar
50	4.0	3.59	2.07
100	11.2	7.18	4.14
150	14.9	10.77	6.21
200	22.4	14.36	8.28
250	26.1	17.95	10.35
300	33.6	21.54	12.42
350	37.3	25.13	14.49
400	44.8	28.72	16.56
450	48.5	32.31	18.63
500	56.0	35.90	20.70

Table 16.2 Inductive reactive power loss of transmission line

Installed capacity/MW	Transmission line loss/Mvar					
	10 km	20 km	30 km	40 km	50 km	60 km
50	0.11	0.21	0.32	0.42	0.53	0.63
100	0.42	0.84	1.27	1.69	2.11	2.53
150	0.95	1.90	2.85	3.80	4.75	5.70
200	1.69	3.38	5.07	6.76	8.44	10.13
250	2.64	5.28	7.92	10.56	13.19	15.83
300	3.80	7.60	11.40	15.20	19.00	22.80
350	5.17	10.34	15.52	20.69	25.86	31.03
400	6.76	13.51	20.27	27.02	33.78	40.53
450	8.55	17.10	25.65	34.20	42.75	51.30
500	10.56	21.11	31.67	42.22	52.78	63.33

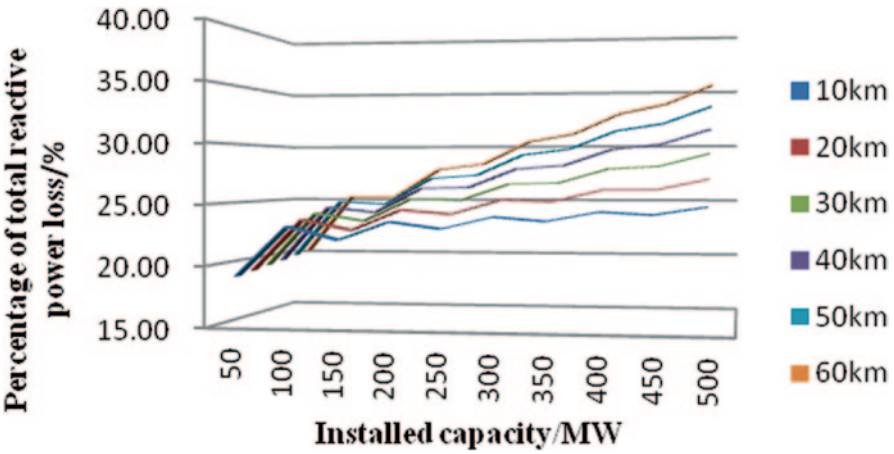


Fig. 16.1 Percentage of total reactive power loss

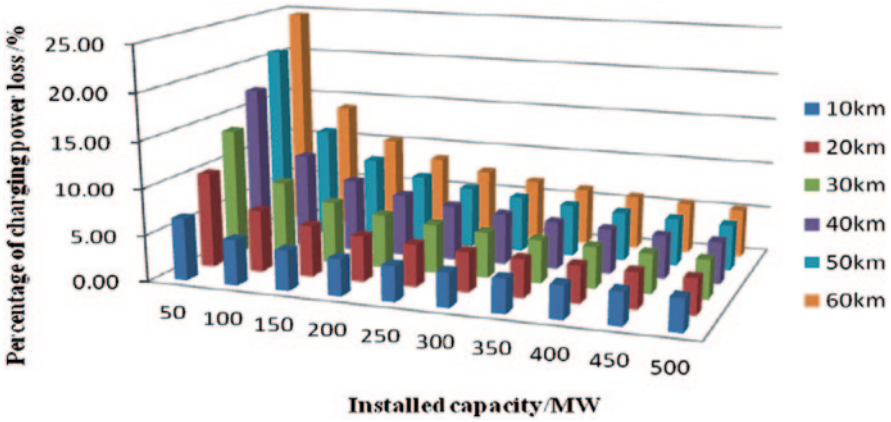


Fig. 16.2 Total charging power loss in percentage of the installed capacity

16.3 Configuration Scheme for Dynamic Device and Capacitor Branches Capacity

16.3.1 Dynamic Reactive Power Compensation Capacity

As for dynamic reactive power compensation capacity, two factors should be considered: first, the coordination with capacitor branches for flexible regulation of reactive power. Taking device regulation ability into consideration [6, 7], the dynamic compensation capacity is proposed not less than 40% of the total compensation capacity.

Second, the dynamic compensation devices' ability to compensate charging power should also be considered. We can calculate the total charging power of different installed capacities (50~500 MW) and different transmission line lengths (10~60 km), as it is shown in Fig. 16.2.

As is shown in Fig. 16.2, the charging power increases with the installed capacity and line length, in which the transmission line length has a greater impact. The installed capacity mainly influences internal charging power of wind farm, and it ranges a little. Although the total charging power increases, le, installed capacity less than or equal to 50 MW and transmission line length greater than or equal to 20 km, installed capacity less than or equal to 100 MW and transmission line length greater than or equal to 40 km, installed capacity less than or equal to 150 MW and transmission line length greater than or equal to 50 km, the total charging power accounts for a larger proportion.

Synthesizing these two factors, it is proposed that dynamic compensation device capacity should not be less than 40% of the total compensation capacity. But in the small-capacity and long lines (such as the above large proportion cases), the total

charging power is likely to exceed 40% of the total compensation capacity. Then, a detailed calculation of charging power is required, as a necessary selection basis for dynamic device.

16.3.2 Capacitor Branches Capacity

After the total capacity of capacitor branches is determined, a distribution process is needed, that is grouping capacitor branches. According to the standards, the maximum capacity of each group of capacitor should not be greater than 8 Mvar in 10 kV voltage level, not be greater than 12 Mvar in 35 kV voltage level, and not be greater than 20 Mvar in 110 (66) kV voltage level.

As wind farm capacitor branches are usually connected to 35 kV bus, the single group capacity should not exceed 12 Mvar. It is inadvisable to set too large single group capacity, and a proper grouping is needed from the perspective of voltage and reactive power control flexibility, whereas investment will be increased for too many groups and more frequently capacitor switching may occur. Based on the two aspects above, capacitor grouping can be determined by the following categories:

1. Dividing 1–2 groups when the total capacity is less than or equal to 12 Mvar.
2. Dividing 2–4 groups when the total capacity is greater than 12 Mvar and less than or equal to 40 Mvar.
3. Dividing more groups when the total capacity is greater.

Then, grouping by an equal capacity according to the actual situation of wind farm.

In order to reduce the inrush current and harmonic level, capacitor branch is generally cooperated with serial reactance. The rated reactance ratio is based on the system harmonic level. Generally, the interference of the 5th and 7th harmonics is serious in wind farm, with which the harmonic current is large. Therefore, series reactance ratio can be chosen between 4.5 and 6%. As for the 3rd harmonic prominent wind farms, a 12% series reactance ratio should be considered.

16.4 Conclusion

Reactive power compensation capacity configuration for wind farm is researched in this paper. The total reactive power loss curve can be easily drawn through the proposed reactive power compensation capacity calculation method, from which we can obtain the needed compensation capacity easily. No less than 40% of the total compensation capacity should be the dynamic compensation capacity, but a detailed calculation of charging power must be considered for the small-capacity wind farm and long lines. Also, a proposal of the smaller than 12Mvar single group capacity capacitor branch and the group numbers is given.

References

1. Bai R, Qin R, Zhi Y et al. Parameter test and analysis for dynamic compensation devices in wind power farms[J]. Electr Power. 2012;45(2):54–7.
2. Yu Y. The research on active disturbance rejection control strategy for SVC in wind farm[D]. Tianjin: Tianjin University of Technology; 2012, p. 3
3. Gu W, Li X-Y, Wei W. Simulation comparison of the improved wind power dynamics by means of SVC and STATCOM[J]. Power Syst Clean Energy. 2009;25(11):70–5.
4. Q/GDW392-2009. Technical rule for connecting wind farm to power grid[S].State Grid. 2010.
5. GB/T19963-2011. Technical rule for connecting wind farm to power system[S]. 2011.
6. Q/GDW212-2008. Technical rule for power system reactive power compensation configuration[S]. State Grid. 2008.
7. Q/GDW 392-2009. Technical rule for connecting wind farm to power grid[S]. 2008; North China Grid.
8. Zhang Q, Qiu W. Design of centralized reactive power compensation capacity of the wind farm [J]. Electr Power. 2012;45(9):25–8.
9. Wang K, Tan Y. Optimized selection of the dynamic reactive power compensation capacity in the wind-power farm [J]. Electr Mach Technol. 2012;5:28–30.
10. Zhang Y, Sun A et al. Reactive power compensation capacity configuration and optimal operation in wind farm [J]. Proc CSU-EPSA. 2011;23(6):150–6.

Chapter 17
External Evaluation of Power Supply Service Based on Zadeh Fuzzy Process

Wenbo Shi, Tianfu Shen and Jian Wang

Abstract To improve the service quality of power enterprise, an index system for power supply service evaluation is constructed from the aspects of electricity customers' perception and the service process of power supply enterprise. G1 method determines the weight of indexes Zadeh operator is used for synthesis process. This chapter combined Zadeh operator and the fuzzy synthesis, and a new evaluation method was proposed. Using this method to evaluate the power supply service at one of the regions of Guangdong Province in China, the result shows the disparity between the actual service performance and customer demands. Through the case, the method of improving power supply service is found out and the reasonableness of proposed method is demonstrated.

Keywords Power supply enterprise · Service quality · External evaluation · Fuzzy evaluation

17.1 Introduction

In the electricity market, grid companies are service-oriented and always operating power supply products. Scientific and rational evaluation of electric power supply service becomes critical to improve service quality [1]. Currently, the fuzzy mode is widely used to evaluate the service quality.

W. Shi (✉) · T. Shen · J. Wang
Electric Power of South China University of Technology, 510641 GuangZhou, China
e-mail: bowenbobo@foxmail.com

T. Shen
e-mail: 455212975@qq.com

J. Wang
e-mail: 78474665@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_17

17.2 Build System of External Evaluation Index

17.2.1 Dimensions of Index

At present, dimension of service quality evaluation that is widely accepted and used internationally is proposed by Parasuraman, Zeithaml, and Berry based on the user perception [2]. The five dimensions of service quality evaluation system include type, reliability, assurance, impact resistance, and empathy [3, 4]. Because of the characteristics of the power industry, we increase two dimensions: the security and stability. This chapter establishes an external evaluation system based on seven dimensions:

1. Type: Service can be observed in the process.
2. Reliability: Reliability refers to the completion of service commitment accurate and reliable, such as the restoration of electricity.
3. Assurance: Assurance is the ability to increase customer confidence in the quality of service and security.
4. Response: It is the ability to make accurate response of the requirements from customers, such as blackout notification timeliness.
5. Empathy: Empathy refers that the power companies according to the customer needs provide differentiated services.
6. Security: Guarantee security of electricity supply companies is important. The power supply enterprise monitors and guides the safe use of electricity.
7. Stability: Stability is the impact of voltage stability and frequency stability.

17.2.2 Index Building

According to the power industry's seven dimensions of service quality evaluation above, based on the feature of power supply enterprise services, combined with supply services at regional power companies in Guangdong Province, we determine the index system for this external evaluation. External evaluation index system in this chapter has set up dozens of secondary indexes. Below the secondary indexes, the third indexes include 40 evaluation components. The names of the secondary and the third indexes are shown in Fig. 17.1.

17.3 Power Supply Service External Evaluation Method

17.3.1 G1 Method Determines Weights

AHP is one of the methods which widely used for weighing purposes. When a large number of elements exist, it is computational intensive and difficult to meet the

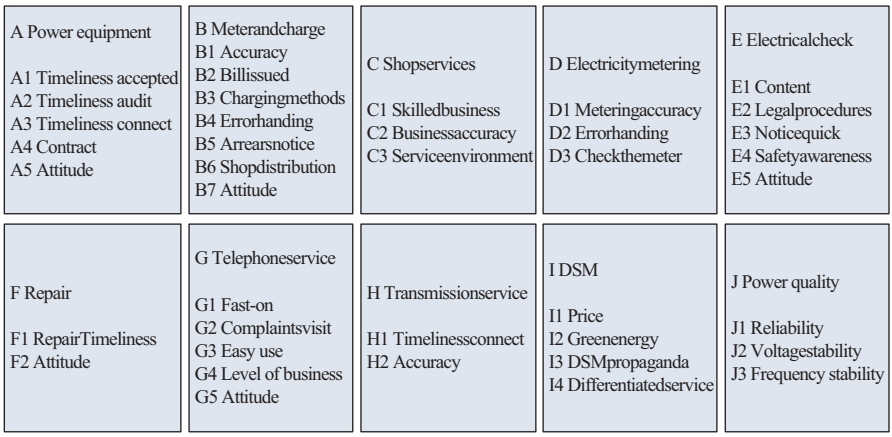


Fig. 17.1 Index system of external evaluation

accurate requirements. G1 method to determine the weights is an improved method of AHP, without judgment matrix, and there is no restriction on the number of elements on the same level [5]. G1 method to determine the weights is as follows [6]:

1. Indexes sorted by importance. Select the most important indicator from the index set and mark I_i, and then select the most important indicator from the remaining indicators and mark I_j. Followed the rule, suppose that a total of m indicators in index set, the indicator sequence relationship is $I_i > I_j > \ldots > I_k$ $(1 < i, j, k < m)$.
2. Determine the relative importance between adjacent indicators. The importance ratio between adjacent indicators is $r_k = \omega_{k-1}/\omega_k$. ω_k is the weight of the k indicators [7]. According to the sequence relationship in the first step, determine the relative importance r_k of each indicator. The relative importance r_k is firstly determined by the experts and then averaged. Relative importance's valuing method is as shown in Table 17.1.

3. Calculate the index weight. The formula is

$$\omega_m = \left[1 + \sum_{k=2}^{m} \prod_{i=k}^{m} r_i\right]^{-1} \tag{17.1}$$

Table 17.1 The value method of r_k

Relationship	r_k value
ω_{k-1} and ω_k equally important	1.0
ω_{k-1} slightly more important than ω_k	1.2
ω_{k-1} obviously more important than ω_k	1.4
ω_{k-1} very more important than ω_k	1.6
ω_{k-1} extremely more important than ω_k	1.8

$$\omega_{k-1} = r_k \omega_k \tag{17.2}$$

$$k = m, m-1, \cdots, 3, 2$$

17.3.2 Evaluation Method Based on Zadeh Operator

Fuzzy evaluation method based on Zadeh operator is as follows:

1. Determine the evaluation factor set $U = \{u_1, u_2, \cdots, u_m\}$, and determine the evaluation factor system.
2. Determine the component set $Q^i = \{q_j^i, (j = 1, 2, \cdots, n)\}$ for each u_i, and n is the number of components.
3. Determine the comment set $V_j^i = \{v_1, v_2, \cdots, v_s\}$ for each q_j^i, and s is the number of comments.
4. Evaluation of the individual indicators. r_{ij} of each component v_i composite an evaluation vector $\alpha_j^i = \{r_{i1}, r_{i2}, \cdots, r_{is}\}$.
5. Establish a fuzzy relationship matrix of centralized component factors:

$$R_i = \begin{bmatrix} \alpha_1^i \\ \alpha_2^i \\ \vdots \\ \alpha_n^i \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1s} \\ r_{21} & r_{22} & \cdots & r_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{ns} \end{bmatrix} \tag{17.3}$$

6. Determine the weight set $\omega^i = \{\omega_1^i, \omega_2^i, \ldots \omega_n^i\}$ of fuzzy evaluation index Q^i, which $\omega_j^i (j = 1, 2, \cdots, n)$ is a fuzzy index weight of $q_j^i (j = 1, 2, \cdots, n)$. Weight set of fuzzy evaluation of external power company is determined by G1 method from the above.
7. Use Zadeh operator and fuzzy synthesis. Fuzzy synthesis process is as follows:

$$B_i = \omega^i \circ R_i = \{b_j^i (1 \le j \le s, j \in \mathrm{N})\} \tag{17.4}$$

8. Using 7-point Likert scale to evaluate the quality of services. Let the service quality level be variable j, and then the 7-point Likert scale evaluation method is as shown in Table 17.2.

Table 17.2 The value of 7-point Likert scale

Comment	Excellent	Very good	Good	Normal	Bad	Very bad	Poor
Level	7	6	5	4	3	2	1

The evaluation of power quality services is A. The formula is as follows:

$$A = \sum_{j=1}^{s} b_j^i \cdot j / \sum_{j=1}^{s} b_j^i \tag{17.5}$$

17.4 Case study

17.4.1 Data Statistics

This chapter assesses an area of Guangdong Province. We design the questionnaire based on the customer's perception of power services. Comment uses 7-point Likert scale. During the calculation, the numbers of each comment questionnaire are divided by the total number of valid questionnaires and obtain the proportion.

This survey selects users in the area including 200 residents and 100 non-residents as a sample. We have given 300 questionnaires and received back 292 questionnaires, in which 268 questionnaires were valid. The effective rate was 89.3%.

Take the third indexes *B2 bills issued timeliness* as an example, and the result of questionnaires is shown in Table 17.3.

Copies of each comment of *B2* are divided by the total number of valid questionnaires' copies. The rating of *B2* is shown in the second line of the matrix *B*.

Statistical analysis gets 10 membership matrixes of 10 secondary indexes. *A* and *B* membership matrixes are listed below, $C \sim J$ abbreviate.

$$R_1 = \begin{bmatrix} 0.135 & 0.108 & 0.245 & 0.209 & 0.177 & 0.061 & 0.065 \\ 0.097 & 0.232 & 0.144 & 0.307 & 0.152 & 0.022 & 0.046 \\ 0.056 & 0.281 & 0.189 & 0.291 & 0.083 & 0.088 & 0.012 \\ 0.173 & 0.198 & 0.204 & 0.368 & 0.028 & 0.021 & 0.008 \\ 0.164 & 0.243 & 0.209 & 0.257 & 0.083 & 0.01 & 0.034 \end{bmatrix} \tag{17.6}$$

$$R_2 = \begin{bmatrix} 0.152 & 0.183 & 0.21 & 0.184 & 0.097 & 0.086 & 0.088 \\ 0.131 & 0.183 & 0.299 & 0.264 & 0.082 & 0.026 & 0.015 \\ 0.021 & 0.102 & 0.081 & 0.368 & 0.151 & 0.205 & 0.072 \\ 0.082 & 0.106 & 0.403 & 0.235 & 0.076 & 0.059 & 0.039 \\ 0.119 & 0.135 & 0.374 & 0.211 & 0.049 & 0.108 & 0.004 \\ 0.067 & 0.158 & 0.105 & 0.423 & 0.108 & 0.112 & 0.027 \\ 0.041 & 0.085 & 0.245 & 0.481 & 0.061 & 0.032 & 0.055 \end{bmatrix} \tag{17.7}$$

Table 17.3 Index *B2* questionnaire results

Comment	Excellent	Very good	Good	Normal	Bad	Very bad	Poor
Copies	35	49	80	71	22	7	4

17.4.2 Weights' Calculation

Fuzzy weights are determined by G1 method. In this assessment, the experts' group determines that the sequence relationship is $J>F>D>I>C>E>B>A>G>H$

Each weight is marked as follows: $\omega_1>\omega_2>\omega_3>\omega_4>\omega_5>\omega_6>\omega_7>\omega_8>\omega_9>\omega_{10}$

According to the weight principles in Table 17.1, the result is as follows: r_2 $r_2=1.2$, $r_3=1.4$, $r_4=1.4$, $r_5=1.2$, $r_6=1.6$, $r_7=1.4$, $r_8=1.4$, $r_9=1.2$, $r_{10}=1.2$ According to the formula (17.1),

$$\omega_{10}=\left[1+\sum_{k=2}^{10}\prod_{i=k}^{10}r_i\right]^{-1}=0.2 \tag{17.8}$$

The rest of the weight is determined by the formula (17.2).

The way of calculating the third indexes' weight is the same as below. The weight of each index is shown in Table 17.4.

Table 17.4 Fuzzy weights of the external evaluation index

Sec	Wet	3th	Wet	Sec	Wet	3th	Wet
A	0.029	A1	0.274	D	0.152	D1	0.463
		A2	0.119			D2	0.331
		A3	0.228			D3	0.207
		A4	0.19	E	0.056	E1	0.177
		A5	0.19			E2	0.337
B	0.04	B1	0.29			E3	0.262
		B2	0.173			E4	0.126
		B3	0.073			E5	0.09
		B4	0.088	F	0.212	F1	0.545
		B5	0.207			F2	0.455
		B6	0.123	G	0.024	G1	0.274
		B7	0.046			G2	0.228
C	0.09	C1	0.276			G3	0.19
		C2	0.255			G4	0.19
		C3	0.28			G5	0.119
		C4	0.189	H	0.108	H1	0.5
I	0.02	I1	0.197			H2	0.5
		I2	0.197	J	0.255	J1	0.488
		I3	0.276			J2	0.265
		I4	0.331			J3	0.265

17.4.3 Evaluation Calculates

According to the formula (17.3), the membership matrix and the three indexes' weight were synthesized. Take the secondary index A as an example.

$$B_1 = \omega_A \circ R_1 = \begin{bmatrix} 0.274 & 0.119 & 0.228 & 0.19 & 0.19 \end{bmatrix} \circ \begin{bmatrix} 0.135 & 0.108 & 0.245 & 0.209 & 0.177 & 0.061 & 0.065 \\ 0.097 & 0.232 & 0.144 & 0.307 & 0.152 & 0.022 & 0.046 \\ 0.056 & 0.281 & 0.189 & 0.291 & 0.083 & 0.088 & 0.012 \\ 0.173 & 0.198 & 0.204 & 0.368 & 0.028 & 0.021 & 0.008 \\ 0.164 & 0.243 & 0.209 & 0.257 & 0.083 & 0.01 & 0.034 \end{bmatrix} = \begin{bmatrix} 0.135 \\ 0.228 \\ 0.245 \\ 0.228 \\ 0.177 \\ 0.088 \\ 0.065 \end{bmatrix}^T \tag{17.9}$$

The rest is similar. The results above constitute the secondary index membership matrix $R_{10\times7}$, the secondary index weight and $R_{10\times7}$ synthesis, the result is

$$B = \omega \circ R_{10\times7} = \begin{bmatrix} 0.126 & 0.212 & 0.255 & 0.255 & 0.106 & 0.09 & 0.09 \end{bmatrix} \tag{17.10}$$

According to the formula (17.4), we obtained the external evaluation of power supply service in this area:

$$\begin{aligned} A &= \frac{\sum_{j=1}^{7} b_j^i \cdot j}{\sum_{j=1}^{7} b_j^i} \\ &= \frac{0.126\times7 + 0.212\times6 + 0.255\times5 + 0.255\times4 + 0.106\times3 + 0.09\times2 + 0.09\times1}{0.126 + 0.212 + 0.255 + 0.255 + 0.106 + 0.09 + 0.09} \\ &= 4.442 \end{aligned} \tag{17.11}$$

External evaluation results of the secondary indexes are shown in Table 17.5.

Table 17.5 Results of the external evaluation

Index	A	B	C	D	E
Point	4.479	4.411	4.282	4.412	4.514
Index	F	G	H	I	J
Point	4.768	4.594	4.821	3.99	4.382

17.4.4 Analyze Evaluation Results

External evaluation result in the area is 4.442 points, which is between the normal and the good. Studies show that the power supply service in this area meets the eligibility criteria ($\geqq$4.2 points). However, there still exists a gap in customers' expectations, so the power company must improve its service in the future. From the results, we can see that on traditional services, such as power transmission services, fault repair, and electrical inspection, the points tend to be better, but in the *C* and *I* indexes, the service quality is inadequate. In times of power market, the power company in this region should further improve service levels in the aspect of direct contacting customers and strengthen the demand-side management.

17.5 Conclusion

In this chapter, the external quality evaluation system for power supply service fully reflects the status of enterprise services. Through the case study, the feasibility of this method is verified. This evaluation method focuses on the role of the main factors, reduces subjectivity, and shows good operability and accuracy. At present, China's electricity market reforms have been carried out. The model in this paper provides a valid reference for the power companies to assess the quality of services under electricity market environment and helps power companies to increase market competitiveness.

References

1. Wang H, Zeng M. Comprehensive evaluation model for power supply service quality based on fuzzy analytic hierarchy process. Power Syst Technol. 2006;30(17):92–5 (In Chinese).
2. Zhou W, Liu L, Chen H. External evaluation of electric power supply service quality based on service blueprint technology. Power DSM. 2009; 11(2):30–3 (In Chinese).
3. Li X, Yang S, Zhang H. Research on assessment method for power supply service quality. Power Syst Technol. 2004;28(12):34–8 (In Chinese).
4. Chi Y, Niu D, Li X. Research on evaluation of power supply service quality based on the matter-element analysis theory. Autom Electr Power Syst. 2010; 34(13):33–7 (In Chinese).
5. Jiang Z. Fuzzy math tutorial. Beijing: National Defense Industry Press; 2004. p. 23–93 (In Chinese).
6. Ren F, Xing M, Niu D. Analysis of matter-element model on evaluation of the reliability of electricity distribution system. Intell Inform Manag Syst Technol. 2005;1(2):312–319 (In Chinese).
7. Zeithaml V. Services marketing integrating customer focus across the firm. Beijing: China Machine Press; 2011. p. 76–97.

Chapter 18
Based on the Power Factors of DFIG Wind Farm for Power Flow Optimization

Xiaohong Hao, Chen Jiang, Li Wu and Lei Zhang

Abstract Double fed induction generators are able to operate on different power factors, and can be used to support power systems. On the basis of regulating the power factors of double fed induction generators, the wind power penetration conditions are taken into account; the hybrid artificial bee colony algorithm is proposed to calculate the optimal power flow, and the impact of different power factors on the power system is researched. In this chapter, the model of objective function is formulated to minimize the conventional generator cost under the consideration of various power factors of double fed induction generators. The simulation results show in high wind speed, double fed induction generators can regulate power factors to support power system operations and reduce the conventional generator cost, and the hybrid artificial bee colony algorithm is better for the converged speed of algorithm than the artificial bee colony algorithm. Thus, the model and new algorithm are proved effectively.

Keywords Double fed induction generator · Power factor · Optimal power flow · Hybrid artificial bee colony algorithm

18.1 Introduction

Recently, more and more people begin to care about the issue of wind power development. What more is that the new wind power technologies play an important role in the wind farm construction. Double fed induction generators (DFIGs) are becoming the mainstream wind turbines for excellent performance and work efficiency [1, 2]. In the power system operation, the large-scale construction of wind farms has growing tendency so that the new type and high install capacity is choosen for wind turbines, which leads to a high wind power penetration. According to this wind power penetration condition, the power systems containing wind power should be

C. Jiang (✉) · X. Hao · L. z. h. Wu · L. Zhang
College of Computer and Communcation, Lanzhou University of Technology,
730050 Lan zhou, China
e-mail: jc_bjnj@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_18

researched to improve the economical efficiency of power systems. Therefore, the optimal power flow (OPF) of a power system containing wind farms is becoming an important issue.

The performance of the DFIG is better than the wind turbine with an asynchronous generator. Generally speaking, the DFIGs operate on the model of the constant power factor, and the power factor is set to 1. However, the worldwide mainstream DFIGs are able to regulate the power factor from 0.95 (lagging) to 0.95 (leading) [3, 4]. Utilizing the regulating capability of DFIG power factor, the DFIGs have flexible power output in the power system. Some researchers have analyzed the impact of DFIG power factors on the power system [4, 5].

The OPF is one of the significant constituents in the power system analysis [6]. The model of the objective function is formulated to minimize the conventional generator's cost, and the DFIG power factors and wind power penetrations are introduced as special wind power conditions. In this chapter, the artificial bee colony algorithm is introduced. The artificial bee colony algorithm is an excellent algorithm. At the same time, this chapter enhances and improves the algorithm performance by combining it with the relevant optimization strategies. In sum, according to the result and theory analysis it indicates that DFIGs can regulate power factors to support the power system, and the model and new algorithm are verified effectively.

18.2 Double Fed Induction Generator Model

In the DFIG operation, both power output and wind speed correspond with each other. This chapter chooses 1.5 MW DFIG as the subject investigated, and the power characteristic and power characteristic curve of DFIG comes from the correlation study [7].

The DFIG power output is decided by power characteristic and wind speed. For the constant power factor control model, the power factor is set as $\cos\phi$. Wind turbine active power P_{T} is expressed as follows [7]:

$$P_{\mathrm{T}} = P_{\mathrm{s}} + P_{\mathrm{r}} = \frac{r_{\mathrm{r}}(x_{\mathrm{s}} + x_m)^2 (P_{\mathrm{s}}^2 + Q_{\mathrm{s}}^2)}{x_m^2 \left|\dot{U}_{\mathrm{s}}\right|^2} + \frac{2r_{\mathrm{r}}(x_{\mathrm{s}} + x_m)}{x_m^2} Q_{\mathrm{s}} + (1-s)P_{\mathrm{s}} + \frac{r_{\mathrm{r}}\left|\dot{U}_{\mathrm{s}}\right|^2}{x_m^2}, \quad (18.1)$$

where P_{T} is the active power of DFIG fed into power system; P_{s} is the active power of stator winding side; P_{r} is the active power of rotor winding side r_{r} is the rotor resistance x_{s} is the stator reactance; x_m is the excitation reactance; $\dot{U}_{\mathrm{s}}$ is the stator voltage; and Q_{s} is the stator reactive power.

The DFIG slip S is calculated from the wind turbines speed characteristic [7]. Because the converter output and input reactive power is very little, the wind turbines reactive power is the stator side reactive power:

$$Q_{\mathrm{T}} = Q_{\mathrm{s}} = P_{\mathrm{s}} \tan \varphi. \tag{18.2}$$

According to this chapter [7], introducing Eq. (18.2) into Eq. (18.1) leads to:

$$P_{\mathrm{T}} = \frac{r_{\mathrm{r}}(x_{\mathrm{s}} + x_m)^2 P_{\mathrm{s}}^2}{x_m^2 \left|\dot{U}_{\mathrm{s}}\right|^2}(1 + \tan^2 \varphi) + (1 + \frac{2r_{\mathrm{r}}(x_{\mathrm{s}} + x_m)\tan \varphi}{x_m^2} - s)P_{\mathrm{s}} + \frac{r_{\mathrm{r}}\left|\dot{U}_{\mathrm{s}}\right|^2}{x_m^2}. \tag{18.3}$$

According to this chapter [5], Q_{s} is expressed as follows:

$$Q_{\mathrm{s}} = \frac{-a\left|\dot{U}_{\mathrm{s}}\right|^2 + \left|\dot{U}_{\mathrm{s}}\right|\sqrt{c\left|\dot{U}_{\mathrm{s}}\right|^2 + 4bP_{\mathrm{T}}}}{2b}\tan \varphi, \tag{18.4}$$

where $a = 1 + \frac{2r_{\mathrm{r}}(x_{\mathrm{s}} + x_m)\tan \varphi}{x_m^2} - s, b = \frac{r_{\mathrm{r}}(x_{\mathrm{s}} + x_m)^2}{x_m^2}(1 + \tan^2 \varphi)$

$$c = (1 - s)^2 + \frac{4r_{\mathrm{r}}(x_{\mathrm{s}} + x_m)\tan \varphi}{x_m^2}(1 - s) - \frac{4r_{\mathrm{r}}^2(x_{\mathrm{s}} + x_m)^2}{x_m^4}.$$

In this chapter, the all active powers and reactive powers of the wind farm are equal to the total of the active power and reactive power by each DFIG in the wind farm [8]. The DFIGs operate on the constant power factor, and the wind farm integrates into the power system by PQ nodes.

18.3 The Model of Power System Optimal Power Flow Containing a Wind Farm

The objective is to minimize the conventional generator cost to ensure that the power system is good and economic [6, 9]. Thus, the objective function is set as follows:

$$\min F = \sum_{i=1}^{N} F_i(P_{Gi}) = \sum_{i=1}^{N} a_i P_{Gi}^2 + b_i P_{Gi} + c_i, \tag{18.5}$$

where a_i, b_i, and c_i are coal-fired coefficients of conventional generators; P_{Gi} is the active power from generator i.

18.3.1 Constraint Condition

$$P_{Gi} - P_{Di} + P_{Ti} = U_i \sum_{j=1}^{n} U_j (G_{ij} \cos\delta_{ij} + B_{ij} \sin\delta_{ij}), \tag{18.6}$$

$$Q_{Gi} - Q_{Di} + Q_{Ti} = U_i \sum_{j=1}^{n} U_j (G_{ij} \sin\delta_{ij} - B_{ij} \cos\delta_{ij}), \tag{18.7}$$

where P_{Gi}, Q_{Gi} are the generator active and reactive power outputs at the i bus. P_{Di}, Q_{Di} are the load active and reactive powers at the i bus. P_{Ti}, Q_{Ti} are the wind farm injected active and reactive powers at the bus i. U_i, U_j are the voltage at the i and j buses, respectively. G_{ij}, B_{ij} are the conductance and the susceptance between the nodes i and j. δ_{ij} is the phase angle between the nodes i and j.

The inequality constraint conditions of conventional generators:

$$P_{Gi.\min} \le P_{Gi} \le P_{Gi.\max}, \tag{18.8}$$

$$Q_{Gi.\min} \le Q_{Gi} \le Q_{Gi.\max}, \tag{18.9}$$

where $P_{Gi.\min}, P_{Gi.\max}$ are the low and upper of active power outputs at the generator i. $Q_{Gi.\min}, Q_{Gi.\max}$ are the low and upper of reactive power outputs at the generator i.

Other inequality constraint conditions are expressed:

$$T_{ap_i.\min} \le T_{ap_i} \le T_{ap_i.\max}, \tag{18.10}$$

$$Q_{c_i.\min} \le Q_{c_i} \le Q_{c_i.\max}, \tag{18.11}$$

$$U_{i.\min} \le U_i \le U_{i.\max}, \tag{18.12}$$

where $T_{ap_i.\max}$ and $T_{ap_i.\min}$ are the low and upper of transformer's tap changers at the i bus; $Q_{C_i.\max}$ and $Q_{C_i.\min}$ are the low and upper of parallel capacitors at the i bus; and $U_{i.\max}$ and $U_{i.\min}$ are the low and upper of voltages at the i bus.

18.4 Hybrid Artificial Bee Colony Algorithm

Karaboga introduced artificial bee colony algorithm first in 2005. Artificial bee colony algorithm consists of three group bees: employed bees, onlooker bees, and scout bees. The three group bees to search food are simulated [10,11]. The employed bees and onlooker bees according to Eq. (18.13) are used to search food source.

$$v_{ij} = x_{ij} + R_{ij}(x_{ij} - x_{kj}), \tag{18.13}$$

where $k \in \{1,2,\ldots,SN\}, j \in \{1,2,\ldots,D\}$, and j and k are random numbers, $k \neq i$. R_{ij} is a random number between 1 and −1.

The onlooker bees are according to the fitness of food source which are feedbacks of employed bees, and the onlooker bees choose the food source with the probability given in Eq. (18.14), whereas fit_i is the fitness:

$$P_i = \frac{fit_i}{\sum_{i=1}^{SN} fit_i}. \tag{18.14}$$

If the fitness cannot improve within the limit, the fitness will be ignored. The employed bees will become scout bees, and a new food source of the scout bees is produced in equation Eq. (18.15):

$$x_i^j = x_{\min^j} + rand(0,1)(x_{\max^j} - x_{\min^j}). \tag{18.15}$$

18.4.1 Boltzmann Selection Strategy

In this chapter, Boltzmann selection substitutes the roulette wheel selection Eq. (18.14). According to the correlation study, Boltzmann selection equations are as given below[12]:

$$P_i = \frac{\exp(fit_i / T)}{\sum_{i=1}^{SN} \exp(fit_i / T)}, \tag{18.16}$$

$$T = T_0(0.99^{C-1}), \tag{18.17}$$

where T is the temperature, T_0 is the original temperature, and C is the iteration.

18.4.2 Self-Adaptive Strategy

The employed bees and onlooker bees are significant parts of search programs in the artificial bee colony algorithm [13]. R_{ij} is the random number and a part of Eq. (18.13). According to the correlation study, the employed bees are expressed with the following equation [13]:

$$R_{ij} = \begin{cases} \exp\left(-k_1\left(\dfrac{iter}{\max iter}\right)\right) & ,rand < 0.5 \\ -\exp\left(-k_1\left(\dfrac{iter}{\max iter}\right)\right) & ,\text{otherwise} \end{cases}. \tag{18.18}$$

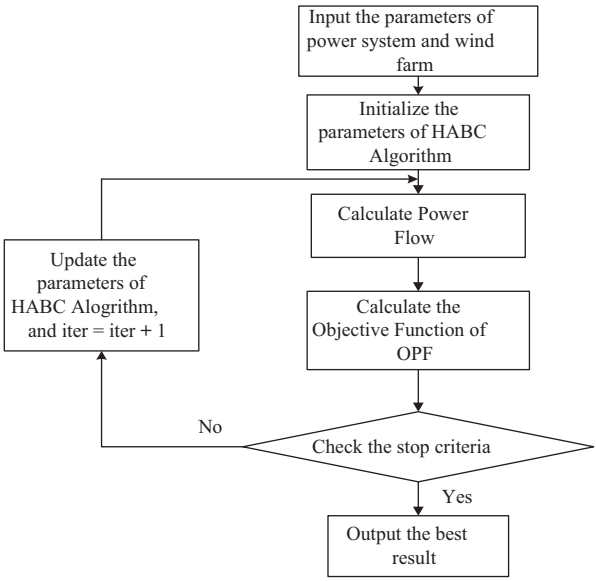


Fig. 18.1 Program flow chart of optimal power flow (*OPF*) by the hybrid artificial bee colony (*HABC*) algorithm

The onlooker bees are expressed with the following equation [13]:

$$R_{ij} = \begin{cases} \exp\left(-k_2\left(\dfrac{iter}{\max iter}\right)\right) & ,rand < 0.5 \\ -\exp\left(-k_2\left(\dfrac{iter}{\max iter}\right)\right) & ,\text{otherwise} \end{cases}, \tag{18.19}$$

where k_1, k_2 are the coefficients; *iter* is the iteration.

According to the correlation study, the program flow chart of OPF by the hybrid artificial bee colony (HABC) algorithm can be expressed as follows [11, 14, 15] (Fig 18.1):

18.5 Simulation Results

In this chapter, the OPF of a wind farm is carried out on the Institute of Electrical and Electronics Engineers (IEEE) 30 bus system in MATLAB [16]. The parameters of conventional generators are taken from the correlation study [9]. The wind farm is connected with the 30 nodes of the system by a transformer. The five parallel capacitors are installed in the 10 and 24 nodes of the system. The transformers are able to adjust to $-4 + 2.5\%$. The parameters of the HABC algorithm and artificial bee colony algorithm have the population size of 50, and the max cycle is 100. In this chapter, the wind speed is mainly considered to be simulated in high wind speed,

Table 18.1 Results of 15% wind power penetration and the power factors 1 and 0.98 (lagging) of three time periods

Time period	Power factor	The voltage of PCC	The power output of conventional generation						Cost (US$/h)
			P1	P2	P5	P8	P11	P13	
1	1	1.0282	166.89	46.45	20.78	15.79	10.35	12.03	731.873
	0.98	1.0397	167.4	46.4	20.78	15.44	10.29	12	731.828
2	1	1.0277	159.18	44.66	20.13	10.59	10	12	679.336
	0.98	1.0468	158.79	44.48	20.16	10.94	10	12	678.806
3	1	1.0235	155.67	43.49	19.97	10	10	12	661.752
	0.98	1.0446	155.37	43.57	19.92	10	10	12	660.898

PCC point of common coupling

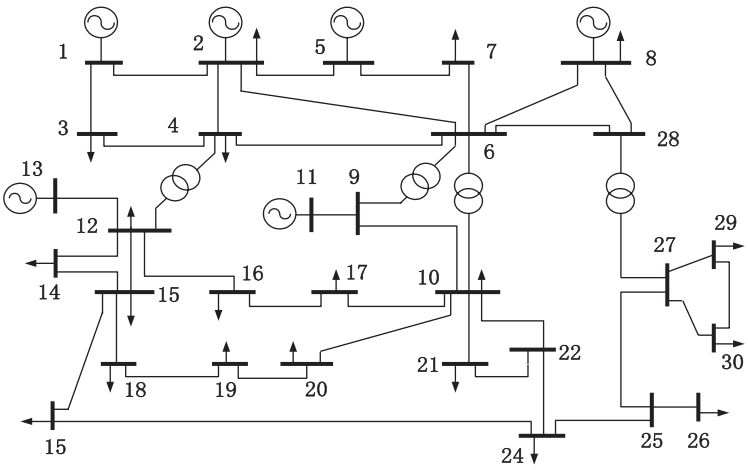


Fig. 18.2 The diagram of IEEE 30 bus test system

the time period 1 is set to 9 m/s. The time period 2 is set to 11 m/s. The time period 3 is set to normal wind speed.

Table 18.1 shows different conditions of wind speed and power factor, and about 15% wind power penetration. In the time period 1, the result of power factor 1 and 0.98 (lagging) has an obvious difference that is the point of common coupling (PCC) voltage, and power factor 0.98 (lagging) can keep PCC voltage at 1.03–1.05. It is the fact that the wind farm of DFIGs can regulate the reactive power output, which is able to support the power system. With the speed becoming normal wind speed, the conventional generator cost is reduced. The difference of the optimization result of the power factors 1 and 0.98 (lagging) is increasing with the wind speed change (Fig 18.2). From the results of the objective function, the data illustrate the operation of the power factor 0.98 (lagging), which is beneficial to the OPF optimization at the high wind speed condition. And the wind farm is able to provide more and more help and support to power systems with wind speed increases.

Table 18.2 Results of 20 and 15% wind power penetration and the power factors 1 and 0.98 (lagging) in normal wind speed

Wind power penetration (%)	Power factor	The voltage of PCC	The power output of conventional generation						Cost (US$/h)
			P1	P2	P5	P8	P11	P13	
15	1	1.0235	155.67	43.49	19.97	10	10	12	661.752
	0.98	1.0446	155.37	43.57	19.92	10	10	12	660.898
20	1	0.9913	146	41.31	19.08	10	10	12	621.398
	0.98	1.0357	144.7	41.23	19.02	10	10	12	616.841

PCC point of common coupling

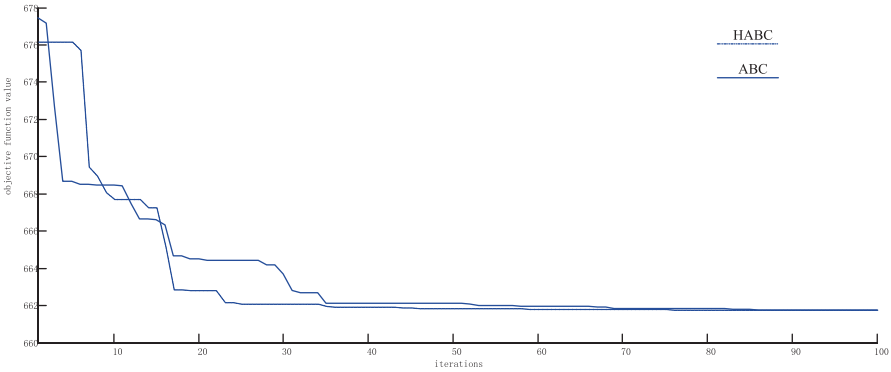


Fig. 18.3 Curve of the hybrid artificial bee colony (*HABC*) algorithm and artificial bee colony algorithm (*ABC*)

Table 18.2 shows the different conditions of power factors between about 20 and 15% wind power penetration in normal wind speed. The comparison between 20 and 15% wind power penetration and power factors 1 and 0.98 (lagging), the Table 18.2 shows that the wind power penetration 20% is the better optimization result than 15%, and the power factor 0.98 (lagging) has better optimization results than the power factor 1. At normal wind speed and 20% wind power penetration, the result of power factor 0.98 (lagging) is reduced to about five conventional generator costs compared with the power factor 1, and the voltage is more stable. From Tables 18.1 and 18.2, the optimization shows that when the conditions are high wind speed and high wind power penetration, the power factor 0.98 (lagging) is beneficial to reduce the power system conventional generator cost and attain OPF.

From Fig. 18.3, the 15% wind power penetration of the DFIG wind farm under the conditions of normal wind speed and power factor 1 is calculated. It can be observed that the HABC algorithm converged speed and is better than the artificial bee colony algorithm. As a result of the HABC algorithm combined with the self-adaptive strategy and Boltzmann selection policy with original algorithm, the iterations of the algorithm are considered in the new algorithm. The new algorithm is verified effectively.

18.6 Conclusion

In this chapter, the OPF under the conditions of DFIG power factors and wind power penetration is researched by the HABC algorithm. In the research, the DFIG power factor 0.98 (lagging) is able to improve the result of the OPF. At high wind speed and wind power penetration, the DFIG power factor 0.98 (lagging) supports the power system that has more reactive power than the power factor 1. From the Fig. 18.3, the simulation results show that the converged speed of the HABC algorithm is fast, and it can draw the conclusion that the new algorithm has an effective improvement. From the Tables 18.1 and 18.2, the simulation results also show that the DFIG power factor 0.98 (lagging) gets better optimization results than using the DFIG power factor 1.

References

1. Lang Y, Zhang X, Xu D. Reactive power analysis and control of doubly fed induction generator wind farm [J]. Proc CSEE. 2007;27(9):77–82 (In Chinese).
2. Tapia A, Tapia G, Ostolaza JX. Reactive power control of wind farms for voltage control applications [J]. Renew Energy. 2004;29(3):377–92.
3. State Grid Corporation of China. Regulations on accessing power system technology to wind electric field in State Grid [R]. Beijing: State Grid Corporation of China; 2009.
4. Mao Q. Compensating reactive power loss of wind farm with reactive power generated by wind turbine generators [J]. Power Syst Technol. 2009;33(19):175–80 (In Chinese).
5. Liu Y, Wang W, Wang Y. Research on wind farm models for power flow calculation [J]. East China Electr Power. 2008;36(4):446–9 (In Chinese).
6. Wang X. Modern power system analysis [M]. Beijing: Science Press; 2003. pp. 116–34.
7. Shen H. Studies on integrated variable-speed constant-frequency wind turbine models and the application [D]. Beijing: China Electric Power Research Institute; 2003.
8. Zhao J, Li X, Hao J. Reactive power control of wind farm made up with doubly fed induction generators in distribution system [J]. Electr Power Syst Res. 2009;80(6):698–706.
9. Abido MA. Multiobjective optimal power flow using strength Pareto evolutionary algorithm [C]. Universities Power Engineering Conference, IEEE; 2004. pp. 457–61.
10. Yue Y. Method of unit commitment based on an improved artificial bees colony algorithm [D]. Zhejiang: Zhejiang University; 2012.
11. Ren X, Zhou L, Zhao F, et al. Reactive power optimization of distribution network based on artificial bee colony algorithm [J]. Mod Electr Power. 2012;29(4):41–5 (In Chinese).
12. Ding H, Feng Q. Artificial bee colony algorithm based on boltzmann selection policy [J]. Comput Eng Appl. 2009;45(31):53–5 (In Chinese).
13. Gu W, Yin M, Wang C. Self adaptive artificial bee colony for global numerical optimization [J]. IERI Procedia. 2012;1:59–65.
14. Sumpavakup C, Srikun I, Chusanapiputt S. A solution to the optimal power flow using artificial bee colony algorithm[C]. 2010 International Conference on Power System Technology, IEEE; 2010. pp. 1–5.
15. Sumpavakup C, Chusanapiput S, Srikun I. A hybrid cultural-based bee colony algorithm for solving the optimal power flow [C]. 2011 IEEE 54th International Midwest Symposium on Date, IEEE; 2011. pp. 1–4.
16. Zhang B, Chen S, Yan Z. Analysis of advanced power network [M]. Beijing: Tsinghua University Press; 2007. pp. 311–3.

Chapter 19
A Novel Method of On-line Measuring and Analyzing the Source Impedance

Wenwu Song, Kai Zhang, Wei Zhang, Mengtao Qin and Bibo Ping

Abstract There are many different ways of measuring the source impedance; however, they vary in applicability and accuracy. The chapter presents a simple and cheap on-line measuring method. It not only applies in AC network but also makes a good result in DC network. The chapter takes emphasis on the common-mode source impedance measuring. Further, the compensation model is established and the calibration factor against frequency is simulated. The simulation results make clear that the measuring frequency range of the method is relatively wide and can reach a good accuracy in coupling analysis and EMI filter design.

Keywords Source impedance · Common-mode EMI · On-line measuring · Power grid noise

19.1 Introduction

Grid of modern society is facing very serious bearer. As considerable progress has been made on the fields applying switching devices of power semiconductor and control technology becomes more sophisticated, the power system may bring out the EMI or shares the EMI of the load during operation. In addition, we know that the power grid supplies to different users. Due to the uncertainty of different users who have access to the electricity, especially the starting and braking of high-power equipment, the neighboring grid will suffer the random momentary shock which may also cause trouble to sensitive users; resulting in abnormal equipment

B. Ping (✉) · W. Song · W. Zhang · M. Qin
Science and Technology on Electromagnetic Compatibility Laboratory, 430064 Wuhan, China
e-mail: qmtqxm@126.com

K. Zhang
China Ship Development and Design Center, 430064 Wuhan, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_19

operation, reducing of the system efficiency, loss of computer data, logic confusion, even hardware damage which leads to a system standstill [1, 2].

In order to effectively estimate the power line interference injecting from different devices and the amount of interference where each electrical equipment coupled from the power line, we need to measure the input impedance (including differential and common mode impedance) of each equipment relatively accurately. It is easy to identify differential mode interference, so this chapter focuses on the measurement and analysis of the common-mode impedance.

The rest of this chapter is organized as follows. Section 19.2 gives the form of the equivalent source impedance. Section 19.3presents the test method in detail. Section 19.4 is the simulation results using the proposed method. And Section 19.5 is the conclusion.

19.2 Form of the Equivalent Source Impedance

For the equivalent input impedance of the power line, it can be divided into common and differential mode impedance. Differential mode impedance is the input impedance of the device. Whichever form of the impedance, it can be expressed as the resistance part and the reactance part. Since we care about the interference of the device as a network coupled to the grid or other devices, therefore, the two parts of the source impedance should be measured and analyzed. The electric model of a network can be expressed as an impedance or an admittance model [3, 4]. The equivalent circuit of the former is shown as Fig. 19.1a and the equivalent circuit of the latter is shown as Fig. 19.1b. Specially, the equivalent impedance (or admittance) of the device is determined at a certain frequency, but the resistance (or conductance) part and reactance (or susceptance) part of the equivalent impedance of different frequencies may be different.

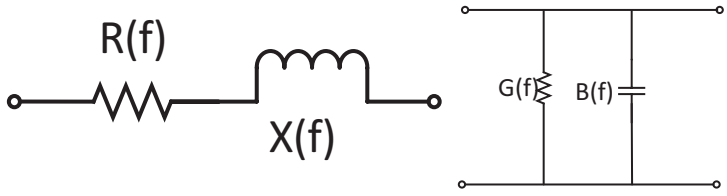


Fig. 19.1 **a** Equivalent circuit in impedance form, **b** equivalent circuit in admittance form

19.3 Test Method

19.3.1 Basic Principles

Ohm's law is the basic principle for the definition, calculation, and measurement of impedance. Schneider proposed the resonance method. The main drawback of this method is that the frequency range proper for measuring is narrow. She proposed a method using a dual current probe which ignores the phase information and the equipment is expensive. Dongbing Zhang proposed the insertion loss method which also ignores the phase information and requires multiple attempts of measurement, and its degree of precision is limited [5, 6].

Impedance of an off-line network is easy to be measured fast and accurately using the network or impedance analyzer. However, these instruments are not applicable if the EUT is on-line. The method we proposed can solve this problem [7].

In order to measure the input impedance of an AC device, the starting frequency when measuring should be at least 10 times the AC grid frequency which can ensure the main harmonics of the grid so that can be avoided. Furthermore, this can protect the instruments and ensure accurate readings.

The main measuring device is the frequency selective voltage meter (FSLM) which must be used in conjunction with a signal generator and an isolation capacitor. This isolation capacitor ought to present a high impedance characteristic responding to grid frequency while making a low impedance access under measured frequency [8]. If the frequency on the line is 50 Hz, the minimum frequency to measure is $10 \times 50\ \text{Hz} = 500\ \text{Hz}$. The test diagram is shown in Fig. 19.2.

The voltage–current measurement method needs to measure the potential of the higher potential side and the output current of the grid. We need a current probe to measure the current. The current probe measures the magnetic field generated by the wire due to the electron motion. In the range specification of the current probe, it can convert the magnetic flux around the conductor to a linear voltage which can be displayed on an oscilloscope or other measuring instruments for analyzing.

After the required voltage, the current values are measured. The absolute value of source impedance and admittance are:

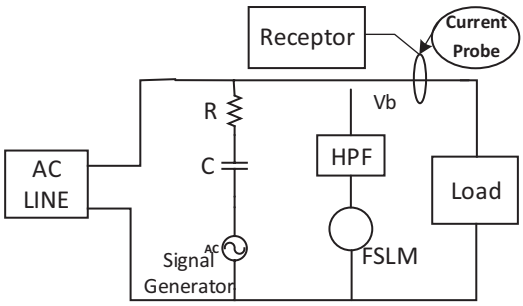


Fig. 19.2 Source impedance testing diagram (voltage–current method)

$$Z(f_0) = \frac{V}{I} \tag{19.1}$$

$$Y(f_0) = \frac{I}{V} \tag{19.2}$$

What we get from the two methods is the modulus value of source impedance or admittance.

19.3.2 On-Line Measurement of the Common Mode Impedance

Differential mode impedance can be measured with the test setup shown in Fig. 19.2. For the common mode impedance, the injection signal and the measurement instruments on the two lines of the grid should be symmetrical. And the signal injected into the line is through the isolation transformer. The setup of the measurement is shown in Fig. 19.3.

a. Test of the common mode impedance or admittance

According to the setup shown in Fig. 19.3, capture the voltage and current at different frequencies. Then the modulus value of the common mode impedance and admittance can be calculated by Eqs. (19.1) and (19.2).

The admittance of an arbitrary network can be expressed as:

$$Y = G + jB \tag{19.3}$$

$$Y^2 = G^2 + B^2 \tag{19.4}$$

where, Y, G, and B are the admittance, conductance, and susceptance of the device at a certain frequency, respectively.

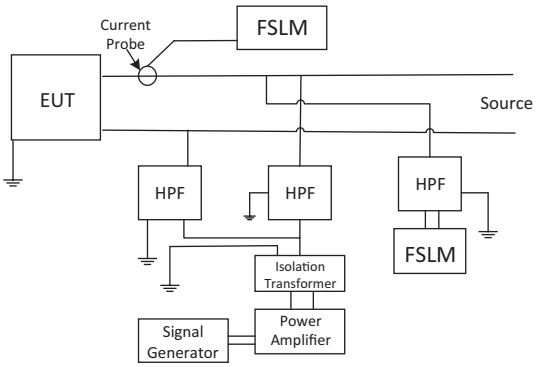


Fig. 19.3 Setup of common mode impedance testing

In order to obtain its resistance (or conductance) part and reactance (or susceptance) part, additional test and proper calculation are necessary.

b. Test when the lines are shunted with capacitors

The additional test measures the common mode admittance when the two lines of the grid are shunted with two equal capacitors. The test method is shown in Fig. 19.3. Supposing the measured admittance is Y_1, so,

$$Y_1 = G + j(B + B_c) \tag{19.5}$$

$$Y_1^2 = G^2 + (B + B_c)^2 \tag{19.6}$$

where B_c is the susceptance of the shunted capacitor.

The common mode admittance test without and with the shunted capacitor establishes the equations as Eqs. (19.5) and (19.7). Solving the two equations and the solutions are as:

$$B = \frac{Y_1^2 - Y^2 - B_c^{\,2}}{2B_c} \tag{19.7}$$

$$G = \sqrt{Y^2 - B^2} \tag{19.8}$$

The corresponding impedance is:

$$Z = \frac{G}{Y^2} - j\frac{B}{Y^2} \tag{19.9}$$

Therefore the resistance (or conductance) part and reactance (or susceptance) part can be obtained. They can also be tested and calculated if two lines of the grid are in series with two equal resistors or inductors. But it is a little bit sophisticated in practical engineering.

c. Selection of the shunted capacitors

The shunted capacitors should make obvious change of the admittance. According to the experience in practical engineering, the admittance of the shunted capacitor should be 1–3 times that of the network at corresponding frequencies.

19.3.3 Design of the High-Pass Filter Cascaded with the FSLM

To ensure the frequency-selective, the voltmeter can withstand the AC voltage of the power grid, a high-pass filter cascaded with the FSLM needs to be designed. A

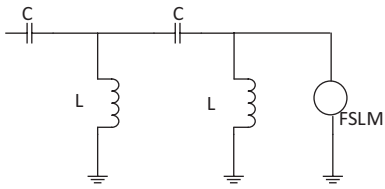


Fig. 19.4 High-pass filter connected with FSLM. *FSLM* frequency selective voltage meter

serial capacitor may be used as the test pins in the form of a probe so as to protect the input terminal of the frequency-selective voltmeter. The cutoff frequency of the high-pass filter must be at least half of the lowest frequency which to be measured.

This filter does not have to be designed specially or can be designed as the Butterworth filter allowing ups and downs in its pass-band [9, 10]. Providing the grid frequency as 50 Hz, the minimum measurement frequency is 500 Hz which requires that the cutoff frequency of high-pass filter is at least 250 Hz. Take for example, 250 Hz, use the K-type high-pass filter which consists of two cascaded identical Γ-shaped high-pass network, as shown in Fig. 19.4:

Since $Z_{FSLM} = 50\Omega$, then we select $K = Z_{FSLM}$, according to the definition of K-type filter, we know

$$L / C = K^2 \tag{19.10}$$

Because the cutoff frequency of the filter is $f_0 = 250Hz$, then

$$\frac{1}{2\pi\sqrt{LC}} = f_0 \tag{19.11}$$

Combine Eqs. (19.3) and (19.4), we obtain

$$L = \frac{Z_{FSLM}}{2\pi f_0} = \frac{50}{2\pi \times 250} mH = 31.8mH \tag{19.12}$$

$$C = \frac{1}{2\pi f_0 Z_{FSLM}} = \frac{1}{2\pi \times 250 \times 50} uF = 12.7uF \tag{19.13}$$

Although the filter is added to protect the test equipment, however, the voltage drop on it cannot be ignored, which needs to be calibrated to compensate the voltage drop.

The voltage calibration factor of the filter is easily obtained by simulation or test, simulation results are shown in Fig. 19.5. We can find that calibration factor is deviated from 1 only in low frequency. When the frequency is high enough, the calibration factor is close to 1 which means the voltage drop of the high-pass filter is small enough comparing with the test voltage [11]. This is consistent with the design result of the high-pass filter.

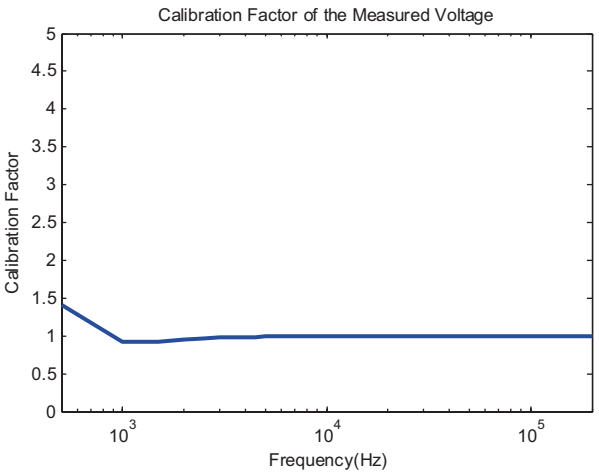


Fig. 19.5 Voltage calibration factor test chart

19.4 Simulation Analysis

19.4.1 Admittance Test

For the common mode admittance simulation, we design a simulation circuit in Fig. 19.6 according to the measurement requirements and the equipment admittance equivalent circuit, the measured voltage is the line-to-ground voltage, so is the current.

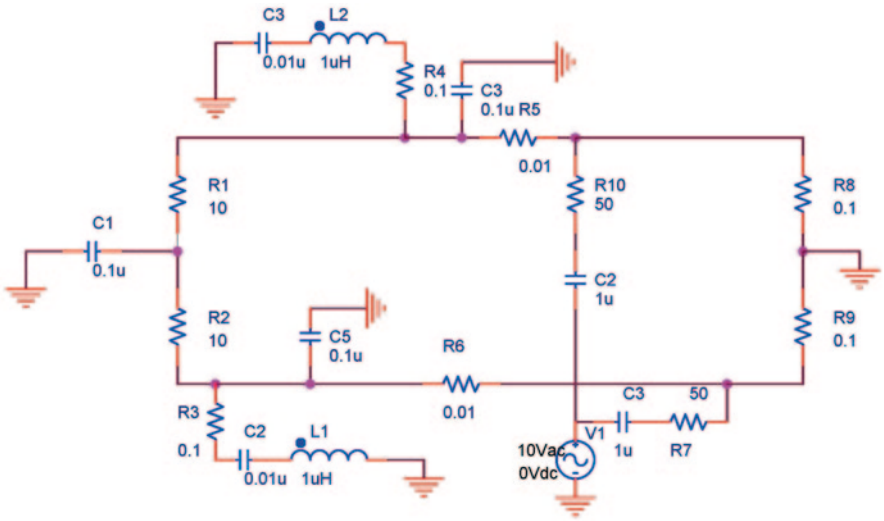


Fig. 19.6 Common mode admittance simulation circuit

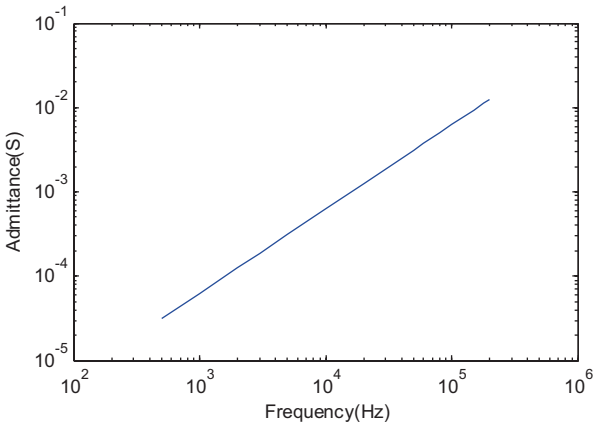


Fig. 19.7 Common mode admittance test results

Wherein, the left side of L1, C2, and R3 constitute the form of impedance to ground equipment. The right side source symmetrically injects the common mode voltage into the grid lines. Use the frequency-selective voltmeter and a high-pass filter to measure the voltage and a current probe to test the current. The ratio between them is an absolute value of the common mode impedance on the equipment side.

After a frequency domain simulation, we can obtain the admittance waveforms against frequency, as shown in Fig. 19.7. Although it is easy to get the modulus values of the admittance for each frequency point one time, the specific form of the equivalent admittance is different under different frequencies.

To calculate the conductance part and the susceptance part of the admittance, a second equation is needed. Therefore, two capacitors can be shunted symmetrically for the two grid lines. Then we can measure the total admittance in the same way. Solving these two equations, we can get the value of the equivalent circuit elements and the conductance part and the susceptance part of the admittance.

19.4.2 Calculation of the Real and Imaginary Part of the Source Admittance

According to the test results of the admittance in Fig. 19.7, we choose a 0.1uF capacitor adding to the two common transmission lines respectively. The admittance test results after adding the capacitors are shown in Fig. 19.8.

The equivalent electrical network before and after shunting the capacitors are shown in Fig. 19.9a, b.

Where $C = 0.1\text{uF}$, $B_c = \text{j} \cdot 2\pi \cdot \text{f}$.

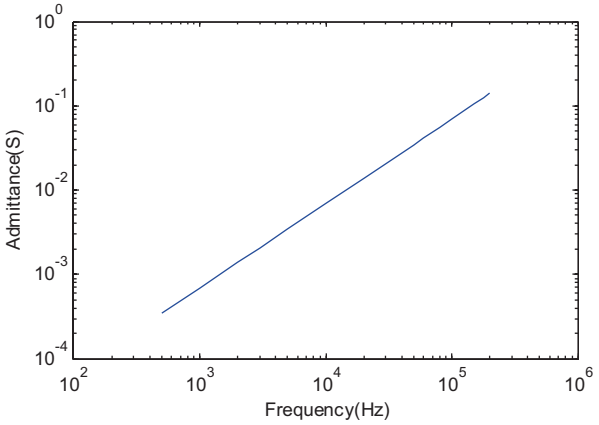


Fig. 19.8 Impedance modulus value after adding capacitors

With Eqs. (19.7) and (19.8), we can obtain the equivalent conductance part and susceptance part of the admittance at different frequencies, as well as the phase information.

The theoretical common mode admittance of the simulation model and the calculation results from Eqs. (19.7) and (19.8) are compared in Table 19.1. The theoretical value refers to the admittance and susceptance of the simulation model we used in PSPISE. And the calculated value refers to the admittance and susceptance we calculated using the proposed method.

According to the results in Table 19.1, the real and imaginary parts of the calculated admittance are in good agreement with the theoretical values, which prove the validity of the test method.

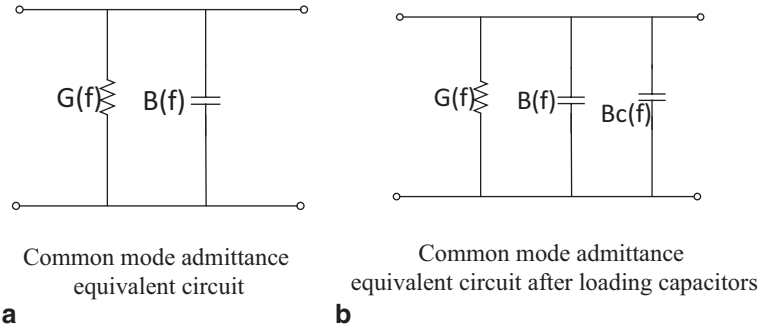


Fig. 19.9 **a.** Common mode admittance equivalent circuit **b.** common mode admittance equivalent circuit after loading capacitors

Table 19.1 Admittance comparison results

Frequency (kHz)	G (theoretical value; S)	G (calculated values; S)	B (theoretical value; S)	B (calculated values; S)
0.5	9.8559×10^{-11}	0.00	3.140×10^{-5}	3.10×10^{-5}
1	3.9438×10^{-10}	0.00	6.280×10^{-5}	6.20×10^{-5}
2	1.5775×10^{-9}	0.00	1.256×10^{-4}	1.26×10^{-4}
3	3.5495×10^{-9}	0.018×10^{-4}	1.8840×10^{-4}	0.0002
5	9.8596×10^{-9}	0.051×10^{-4}	3.140×10^{-4}	0.0003
10	3.9439×10^{-8}	0.202×10^{-4}	6.280×10^{-4}	0.0006
15	8.8738×10^{-8}	0.455×10^{-4}	9.420×10^{-4}	0.0009
20	0.00	0.00	0.0013	0.0013
30	0.00	0.00	0.0019	0.0019
50	0.00	0.00	0.0031	0.0032
60	0.00	0.00	0.0038	0.0038
80	0.00	0.00	0.0050	0.0052
100	0.00	0.00	0.0063	0.0066
150	0.00	0.00	0.0094	0.0105
180	0.00	0.00	0.0113	0.0133
200	0.00	0.00	0.0126	0.0152

19.5 Conclusion

Based on actual engineering, this chapter proposed a simple and convenient source impedance (admittance) test method which can measure the absolute of the impedance and admittance, as well as their real and imaginary parts. The method uses the high-pass filter to protect the monitor instruments from the interference of the AC grid line. The simulation results demonstrate the effectiveness of the test method.

References

1. Chen Y, Qiang Shi. Non-ignorable disturbance source impedance. Saf Electromagn Compat. 2010;12(1):46–57.
2. Zhang D, Chen DY. Measurement of noise source impedance of off-line converters. IEEE Trans Power Electron. 2000;15(5):820–5.
3. Zhang TZ. Resistance network model. Xi'an: Xidian University Press; 2011. p. 15–31.
4. Hua R, Zhang Lei, Hui Long. Analysis of insertion loss of EMI filter. Electron Sci Technol. 2011;24(1):65–7.
5. Liu C. Design of signal EMI filter with impedance mismatch. Electron Meas Technol. 2009;32(7):9–13.
6. Jin M, Ma W. A new technique for modeling and analysis of mixed- mode conducted EMI noise [J]. IEEE Trans Power Electron. 2004;19(6):131–5.
7. Ozenbaugh RL, Pullen TM. EMI filter design. Boca Raton: CRC; 2011. p. 122–210.

8. Zhiyong T, Li J. An improved switch power conducted EMI source impedance test method. Mass electron. 2010;21(4):72–3, 77.
9. Vuttipon T. Accurate extraction of noise source impedance of an SMPS under operating conditions. IEEE Trans Power Electron. 2010;25(1):111–7.
10. Qian W, Han J. The impact of impedance on measurement. Foreign Electron Meas Technol. 2006;25(4):67–9.
11. Wei Y, Yang Z, Xiao-quan L, et al. A modified EMI noise source impedance modeling by employing two resistances calibration and Levenberg-Marquardt's method. 2010 International Conference on Microwave and Millimeter Wave Technology. IEEE, 2010. p. 2021–4.

Chapter 20
Power Flow Calculation for Weakly Meshed Distribution Network with Distributed Generations and Electric Vehicles

Lingyun Wang, Xuanqing Zhou and Yuan Liu

Abstract The conventional power flow algorithms are not applicable to the distribution network with distributed generations (DGs) and electric vehicles because of various node types of generations and weekly meshed network. Based on the analysis of the operation mode and the control characteristics of conventional DGs and electric vehicles, their mathematical models in power flow calculation can be established. Then, an improved power flow calculation method based on the Newton algorithm is presented to compute the power flow of weakly meshed distribution network with multitype DGs and electric vehicle. Finally, an Institute of Electrical and Electronics Engineers (IEEE) 33 bus system is used to validate the proposed algorithm with the results showing that this improved algorithm is feasible and effective.

Keywords DGs · Electric vehicles · Distribution network · Power flow calculation

20.1 Introduction

The traditional distribution network usually uses the form of radiation and the power flow direction is unidirectional. After the distributed generation (DG) and electric vehicle are integrated to the distribution network, the distribution network is converted from the original single generation systems into a multiplicity generations systems, and producing bidirectional power flow, which has exerted enormous

X. Zhou (✉) · L. Wang · Y. Liu
College of Electrical Engineering & New Energy, China Three Gorges University, 443002 Yichang, China
e-mail: 1649082217@qq.com

L. Wang
e-mail: wly@ctgu.edu.cn

Y. Liu
e-mail: 442359141@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_20

impact on the distribution network structure, power loss, voltage profile, and power flow calculation. In order to deal with these issues, an improved power flow algorithm is proposed in this chapter. It is significant to study the distribution network with DGs and electric vehicles.

As to the power flow algorithms of distribution network with the DGs and electric vehicles, Yang et al. applied the Newton method for the power flow calculation; the DGs in distribution network can be considered as PQ and PV nodes simply; but the great errors could not be avoided in this calculation [1]. Wang et al. took into account the output of active power of asynchronous generator and the absorbed reactive power from the distribution network, the P-Q(V) steady-state models of wind turbine were established based on equivalent circuit of induction generator [2]. Li et al. applied back/forward sweep method to calculate the power flow of distribution network with DGs and proposed the reactive power compensation method to deal with the distribution network with PV node type; but meshed network was not considered [3]. The large-scale electric vehicles are connected into a power system, which brings a great deal of influences on power system operation [4]. As the electric vehicles have the characteristics of bidirectional energy flow, the combination of electric vehicles and DGs can overcome the inherent power fluctuations of renewable energy generation and reduce the impact of renewable energy generation on power systems. Fernandez L. P. et al. analyzed the effect of electric vehicles on the distribution network; the electric vehicles were treated as constant power loads [5]. Compared with the back/forward sweep method, the Newton method can be better used to handle more node types directly and meshed network. The weekly meshed network and the mode of electric vehicles are not considered in many existing methods. In this chapter, the given method not only considers the above problems, but also combines both the electric vehicles and DGs. In this chapter, the mathematical models of different typical DGs and electric vehicles are established. Then, an improved power flow calculation method based on the Newton algorithm is presented to compute the power flow of weakly meshed distribution network with multitype DGs and electric vehicles and network. Finally, the algorithm is tested on an IEEE 33 bus distribution system.

20.2 Models of DGs in Power Flow Calculation

20.2.1 Wind Turbines

At present, there are mainly four types of wind turbines in which the constant velocity and the slip-type asynchronous wind turbines can be considered as P-Q(V) nodes in power flow calculation; direct-drive synchronous and doubly fed wind turbines can be considered as PQ nodes [6]. As to the P-Q(V) steady-state models of wind turbine, compared with the asynchronous generator, the difference is that wind turbine is a local compensation of reaction power by means of automatic grouping switch of parallel capacitors; and it does not depend on the power grid to establish magnetic fields. The injected reactive power of node is calculated by $Q = Q'' - Q'$, where $Q''$is the output of reactive power of capacitors, $Q'$is the absorbed reactive power of wind turbines [2].

20.2.2 Photovoltaic Systems

Photovoltaic system can convert solar energy into electricity with their output of direct current (DC) power converted via an inverter to grid compatible alternating current (AC) power. The main category of inverters is current control mode (constant current). The photovoltaic systems can be modeled as PI nodes [7]. The injected reactive power can be written as follows:

$$Q = \mathrm{sqrt}\left(|I|^2 |U|^2 - P^2\right), \tag{20.1}$$

where U is the voltage amplitude of nodes, I is the output of constant current amplitude of photovoltaic systems, and P is the output of constant active power of photovoltaic systems.

20.2.3 Fuel Cells

The fuel cells can convert chemical energy into electricity with their output of DC power converted via an inverter to grid compatible AC power. The main category of inverters is the voltage control mode (constant voltage). The fuel cells can be modeled as the V nodes [8]. The values of injected active and reactive powers can be calculated as follows:

$$P = mU_{\mathrm{FC}}U_S \sin\delta / x \tag{20.2}$$

$$Q = mU_{\mathrm{FC}}U_S \cos\delta / x - U_S / x, \tag{20.3}$$

where x is the equivalent reactance of the transformer that connects the fuel cells and the grid; U_{FC} is the output DC voltage of fuel cells; U_S is the nodes voltage of the system; m is converter modulation depth; and δ is the output AC voltage phase angle of fuel cells. The active and reactive powers of the fuel cells are obtained by controlling the inverter parameters m and δ.

20.3 Electric Vehicles

The core component of electric vehicles is the accumulator. Under the mode of vehicles plug-in with logic (V1G), the electric vehicles can be molded as the power grid load, which can be considered as PQ nodes in the power flow calculation. Under the mode of vehicle to grid (V2G), the electric vehicles can access the power grid to be charged; the accumulator works at the rectifier state and the energy is converted from the power grid side into the DC side of accumulator; in this sense,

the electric vehicles can be considered as a power grid load, which can be considered as PQ nodes in power flow calculation. Under the mode of V2G, the electric vehicles also are capable of injecting energy into the power grid; the accumulator works at the inverter state; and the energy is converted from DC side of accumulator into the power grid side. The conventional charging method of electric vehicle adopts the constant current discharging so that the electric vehicle can be molded as PI nodes in the power flow calculation. The injected reactive power can be written as follows:

$$Q=sqrt(|I|^2 |U|^2 - P^2), \tag{20.4}$$

where U is voltage amplitude of nodes, I is the output of constant current amplitude of electric vehicle, and P is the output of constant active power of electric vehicle.

20.4 Power Flow Algorithm of Distribution Network with DGs and Electric Vehicles

The models of node, branch, transformers, PQ nodes, PV nodes, PI nodes, and P-Q (V) node are established, respectively, in MATLAB. The detailed steps of power flow algorithm of weakly meshed distribution network with DGs and electric vehicles shows below:

1. Input data, DGs, and electric vehicles are divided into the type of P-Q (V), PQ, PI, and PV nodes.
2. Generate the nodal admittance matrix Y, the initial number of iterations $k=0$, precision $\varepsilon=10^{-7}$.
3. P-Q (V), and PI nodes are converted into PQ node, then the injected node power P and Q are calculated.
4. The imbalance power $\Delta P, \Delta Q$ can be calculated according to the Newton method. Judge the formula $\max\{\Delta P, \Delta Q\} < \varepsilon$. If it is satisfied, the calculation ends and the result of the power flow is the output; otherwise, generate Jacobian matrix J; then the voltage modification can be gained based on the correction equation of the Newton method to update the voltage of each node, and then return to Step 3.

20.5 Case Studies

The program flow chart of an improved power flow calculation method based on the Newton algorithm is shown in Fig. 20.1.The IEEE 33 bus system is applied to validate the algorithm in this chapter [6]. The IEEE 33-bus distribution network with DGs and electric vehicles is shown in Fig. 20.2. The standard power is 10 MVA with the standard voltage of 12.66 KV. In order to study the influence of weakly meshed

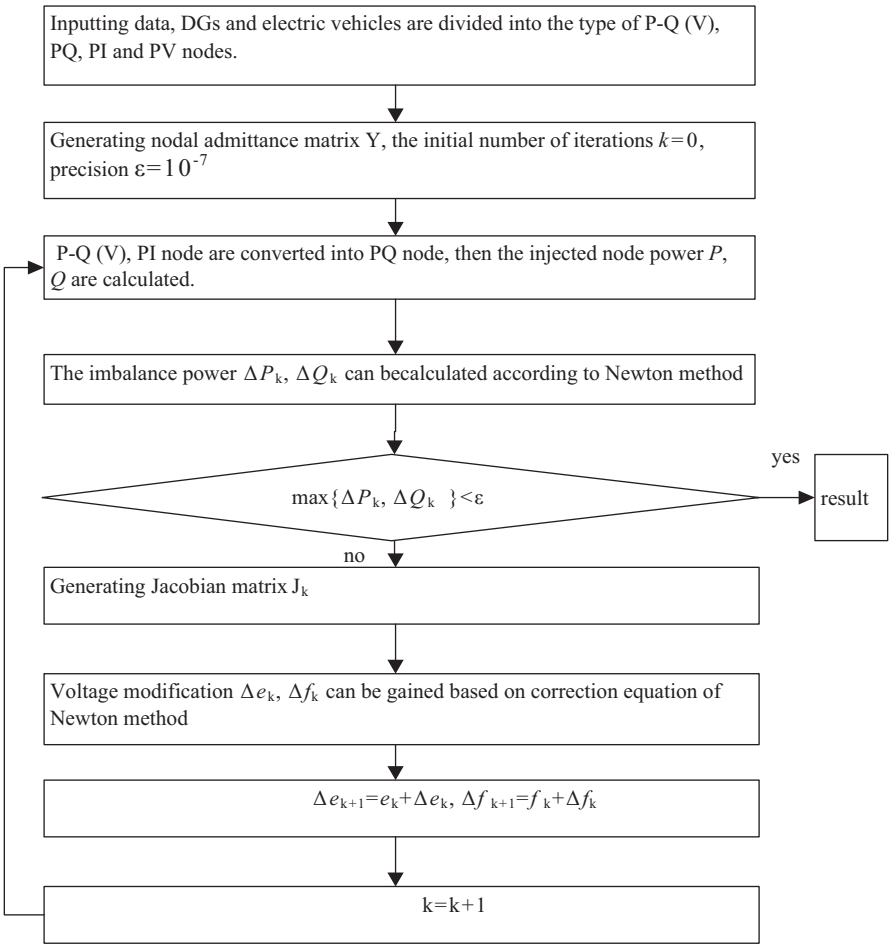


Fig. 20.1 Program flow chart of an improved power flow calculation method

distribution network with DGs and electric vehicles, and feasibility of algorithm, there are nine cases in Table 20.1. The number of iterations and some specific nodes voltage of each case are shown in Table 20.2.

From the comparison of case 1, case 2, case 3, case 5, and case 6 in Table 20.2, it can be seen that when DGs are connected into the distribution network and the nodes voltage amplitude are improved at different degrees because the DGs can inject power into the distribution network, and the voltage increment is relevant to the output power of DGs. Comparing case 3 with case 4, the farther the location of PQ type DG from the root node, the greater the support capacity of the node voltage will be. Comparing case 1 with case 7 and case 8, it can be seen that the change of the node voltage amplitude when electric vehicles are connected into the distribution network because the electric vehicles can be considered as the distribution network loads or generators. When the distribution network loads, its

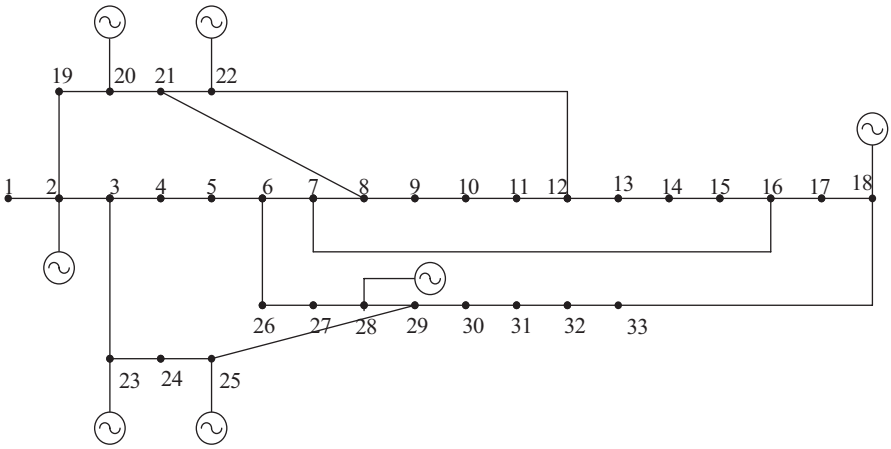


Fig. 20.2 Institute of Electrical and Electronics Engineers (*IEEE*) 33-bus distribution network with distributed generations (DGs) and electric vehicles

Table 20.1 Description of cases

Case	Type of DGs and electric vehicles	Parameters of DGs and electric vehicles	Access
1	None	None	/
2	Wind turbines (P-Q(V) node)	P=300 KW, $\cos\varphi_2=0.98$	22
3	Wind turbines (PQ node)	P=300 KW, Q=200 KVar	20
4	Wind turbines (PQ node)	P=300 KW, Q=200 KVar	2
5	Photovoltaic systems (PI node)	P=150 KW, I=15A	25
6	Full cells (PV node)	P=150 KW, V=12.66 KV	28
7	Electric vehicles (PQ node)	P=300 KW, Q=200 KVar	23
8	Electric vehicles (PI node)	P=100 KW, I=10 A	18
9	Combination of case 2, 3, 5, 6, and 7		

Table 20.2 Number of iterations and nodes voltage of each case

Node	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
1	1	1	1	1	1	1	1	1	1
2	0.9971	0.9973	0.9973	0.9973	0.9972	0.9986	0.9969	0.9972	0.9985
3	0.9862	0.9868	0.9869	0.9865	0.9869	0.9942	0.9848	0.9866	0.993
4	0.9826	0.9832	0.9834	0.9828	0.9834	0.993	0.9811	0.9831	0.9919
6	0.9711	0.9723	0.9726	0.9713	0.9723	0.9912	0.9696	0.972	0.9919
9	0.9656	0.9675	0.968	0.9659	0.9668	0.9835	0.9644	0.967	0.9857
20	0.9807	0.9825	0.9846	0.981	0.9814	0.9901	0.98	0.9814	0.9941
25	0.9626	0.9636	0.9639	0.9629	0.9653	0.9873	0.9605	0.9639	0.9878
28	0.9637	0.9648	0.9651	0.9639	0.9656	1.0001	0.9619	0.965	1.0008
32	0.9532	0.9546	0.9548	0.9534	0.9552	0.9793	0.9514	0.9558	0.9809
33	0.9534	0.9548	0.9551	0.9536	0.9554	0.9788	0.9517	0.9563	0.9802
Iteration	3	4	3	3	4	3	3	4	4

nodes voltage amplitude will be reduced. Like the distribution network generators, its nodes voltage amplitude is improved. Comparing case 8 and case 9, it can conclude that when the DGs are connected into the distribution network, it will reduce the impact of electric vehicles of load type on the distribution network. We can see the number of iteration that does not increase obviously when DGs and electric vehicles are connected into distributed grid from Table 20.2; in this sense, this algorithm can assure the calculation speed. Table 20.2 shows the improved algorithm is feasible and effective to deal with the weakly meshed distribution network with DGs and electric vehicles. In Table 20.1, the solar energy can not only be used in this system; other new energy can be used in this system instead of the solar energy.

20.6 Conclusion

Based on the analysis of operation mode and control characteristics of the wind turbines, the photovoltaic systems, the fuel cells, and the electric vehicles, their models in power flow calculation are established. Considering the meshed network issue, an improved power flow calculation method based on the Newton algorithm is presented to compute the power flow of weakly meshed distribution network with the multitype DGs and electric vehicle. The algorithm has been tested on the IEEE 33 bus distribution systems with the conclusions obtained as below:

1. This algorithm can be used to handle the problem of multitype DGs and meshed network.
2. When DGs are connected into the distribution network, its node voltage amplitude is improved at different degrees; the number of improved voltage is related to the output of power of DGs. The farther the location of DG from the root node, the greater the support capacity of the node voltage will be.
3. The charging/discharging behavior of large-scale electric vehicle will bring great influence on the power grid and the storage characteristics of electric energy will offer new opportunities for the safe and economic operation of the distribution network with DGs.

Acknowledgments This work is supported by the National Natural Science Foundation of China (51407104).

References

1. Yang X, Duan J. Power flow calculation for distribution network with distributed generation. Power Syst Technol. 2009;33(18):140–3 (In Chinese).
2. Wang S, Jiang X. Power flow analysis of distribution network containing wind power generators. Power Syst Technol. 2006;30(21):42–5 (In Chinese).

3. Li X, Peng Y, Zhao J-J. Power flow calculation for distribution network with distributed generation. Power Syst Prot Control. 2009;37(17):78–81 (In Chinese).
4. Divya KC, Ostergaard J. Battery energy storage technology for power systems—an overview. Electr Power Syst Res. 2009;79(4):511–20.
5. Fernandez LP, Roman TGS, Cossent R. Assessment of the impact of plug-in electric vehicles on distribution networks. IEEE Trans Power Syst. 2011;26(1):206–13.
6. Ding M, Guo X. Three-phase power flow for the weakly meshed distribution network with the distributed generation. Proc CSEE. 2009;29(13):35–40 (In Chinese).
7. Naka S, Genji T, Fukuyama Y. Practical equipment models for fast distribution power flow considering interconnection of distributed generators. Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference, IEEE, 2001. pp. 1007–12.
8. El-Sharkh MY, Rahman A, Alam MS. Analysis of active and reactive power control of a stand-alone PEM fuel cell power plant. IEEE Trans Power Syst. 2004;19(4):2022–8.

Chapter 21
Security Monitoring Technology of Smart Grids

Wenjing Li

Abstract In view of the frequent occurrence of current electric power theft accident in combination with existing domestic and foreign research status, a technology research of grid of intelligent security monitoring is proposed in this chapter. It describes the transmission line intelligent video surveillance system, analyzes the target detection algorithm, contrasts Kalman filter background modeling method and improves the target detection method, presents a case study with improved target detection algorithm, and is thus verified by the Halcon. This method has been successfully applied to the transmission line intelligent video surveillance system in Xinjiang Electric Power Company, 220 Kv. It provides a powerful guarantee for the power system security.

Keywords Transmission lines · Intelligent surveillance system · Intelligent detection · Image processing

21.1 Introduction

The power industry is not only the basic industry of strong public good, but also the basic guide industry of national economic development. It is the technology and capital-intensive industry. The safe and stable operation of power systems is related to the reliability of power supply and the sustainable development of national economy. However, as the power facilities feature multipoint, large scope, extensive coverage and bare field, it is hard to maintain the same [1]. Someone even will power facilities as a "material warehouse without walls" to steal stayguy, bolts, Steel-Clad and tower, etc. Besides, with the acceleration of economic construction development, the project constructions are increasing across the country. The towers and electric wires damaged by large machinery pose a great threat to our life [2].

W. Li (✉)
Chongqing Water Resources and Electric Engineering College,
402160 Chongqing, China
e-mail: wawj2013@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_21

A transmission line intelligent video surveillance system based on the differentiation analysis is presented. It is based on visual and pattern recognition. Apart from the traditional monitoring, the system also adds automatic processing and analysis to the image frame extracted from the transmission line video stream, mainly to achieve the following functions: (1) to detect whether there is a large machinery in the surrounding environment of the transmission line, preventing the power accidents caused by large machinery's barbaric construction near the tower or bumping against the tower; (2) to detect whether there is staff activity in the specific area near the transmission line tower, preventing the power accidents caused by pedestrians climbing the tower or stealing the tower material of the transmission line system. It has achieved the function of intelligent monitoring.

21.2 Transmission Line Intelligent Video Surveillance System

The transmission line intelligent video surveillance system proposed in this chapter adopts distributed intelligent system architecture and modular design. It mainly consists of local surveillance terminal subsystem, transmission subsystem, surveillance center management subsystem and intelligent terminal. The local surveillance subsystem is mostly installed in the transmission line tower or in the building near the transformation line. The system is composed of the main control unit which is made up of DSP plus ARM, PTZ video camera, sound and light alarm, video acquisition processing and wind and solar power supply system, etc. It uses the video recognition algorithms and the embedded image-processing technology to complete the live video image's real-time processing, the abnormal state of intelligent recognition, and the input and output of alarm linkage, etc. The transmission subsystem sends the processed signal to the surveillance center management subsystem by wire and wireless way; then, the surveillance center management subsystem sends the diagnostic information and abnormal images to a PC client or smart handheld terminal by SMS/MMS form. The handling terminals, the surveillance center PC client control PTZ and lens are required to complete the storage management and remote communication. The system components are shown in Fig. 21.1.

21.3 Target Detection Algorithms

This chapter focuses on the moving target detection method in a static background environment [3].

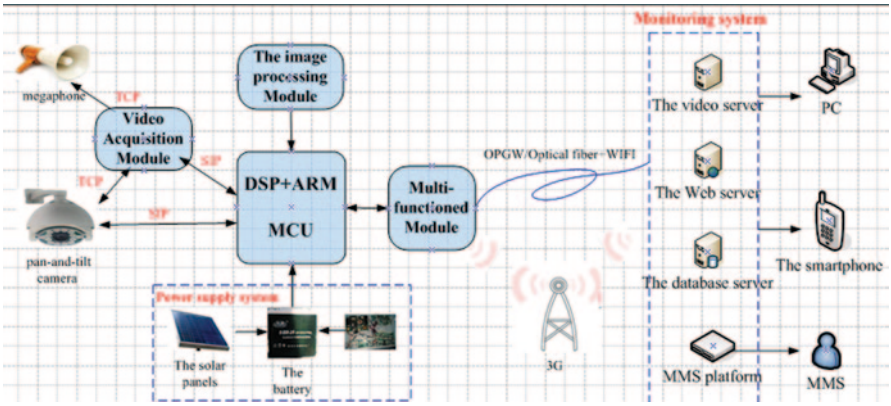


Fig. 21.1 System components schematic diagram

21.3.1 Three-Frame Difference Method

The basic idea of inter-frame difference method [4] is to detect changes in the pixels of adjacent frames in the video image sequence during a very small time interval Δk (usually $\Delta k << 1$ s). When the location of the moving target in the images changes, the gray value of the moving target in the new location in the images will change accordingly and those unchanged objects' gray value in the pixel point of images changes a little. Subtract the adjacent frame images and detect the pixel point whose gray value changes a lot; then, we will get the location and contour of the detected moving target. This chapter studies three-frame difference method [5]. The detection principle is shown in Fig. 21.2.

21.3.2 Kalman Filter Background Modeling Method

Kalman filter is based on the signal in the past by using statistical characteristics of past time varying random signals, based on linear minimum mean-squared error as the optimization criteria, estimating the current or the future value of random signal as close as possible to the true value. It is a time-domain recursive filtering method. Based on the principle of the Kalman filter, it is suggested using Kalman filter to update background model [6, 7]. Set $\hat{s}(x, y, t_k)$ for background model feature of the pixel point (x, y) at time k, $\tilde{s}(x, y, t_k)$ for the background model predictive features of the pixel point (x, y) at time k, $z(x, y, t_k)$ for the measurement characteristics of pixel (x, y) at time k. Based on Kalman filter definition, the conclusions are as follows:

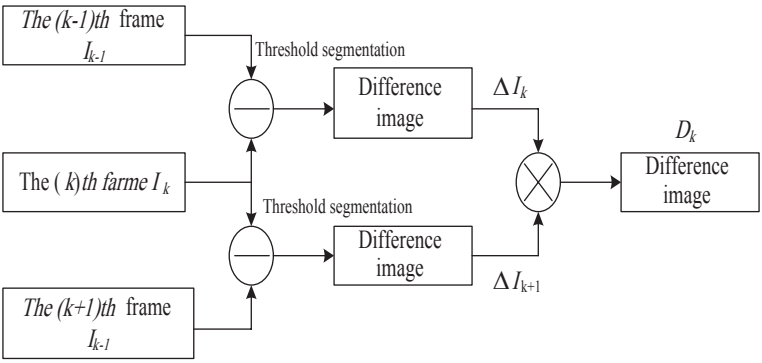


Fig. 21.2 Principle of the three-frame difference method

$$\hat{s}(x,y,t_k) = \tilde{s}(x,y,t_k) + K(x,y,t_k)[z(x,y,t_k) - H(x,y,t_k)\tilde{s}(x,y,t_k)] \quad (21.1)$$

$$\tilde{s}(x,y,t_k) = A(x,y,t_k)\hat{s}(x,y,t_{k-1}) \quad (21.2)$$

where $A(x,y,t_k) = 1$, $H(x,y,t_k) = 1$, the state estimation equations of Kalman filter is shown in formula (21.1), the formula (21.2) is a pixel point in the Kalman filter state prediction equation.

The image is divided into regions containing the moving foreground and regions with relatively static background. To establish a more accurate background model to weight and update foreground and background images respectively, set Kalman again:

$$K(x,y,t_k) = \alpha c(x,y,t_k) + \beta[1 - c(x,y,t_k)] \quad (21.3)$$

$$c(x,y,t_k) = \begin{cases} 1 & d(x,y,t_k) > Th(t_k) \\ 0 & d(x,y,t_k) \le Th(t_k) \end{cases} \quad (21.4)$$

$$d(x,y,t_k) = \left| z(x,y,t_k) - \hat{s}(x,y,t_k) \right| \quad (21.5)$$

where $c(x,y,t_k)$ is the estimated mark of the background model, β is background factor, α is foreground factor, and $Th(t_k)$ is the segmentation threshold to extract foreground target. The α and β values are based on experience, such as taking $\alpha = 0.001$ and $\beta = 0.03$ in this chapter.

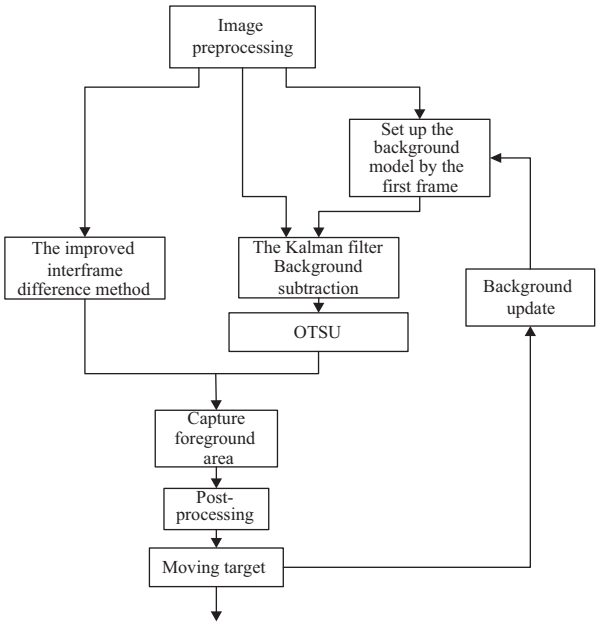


Fig. 21.3 Improved algorithm flow chart

21.3.3 Improved Target Detection Algorithm

A novel real-time target detection algorithm based on combining inter-frame difference and the background subtraction was presented. First, the improved inter-frame difference method is used to detect the moving target as moving target area 1. At the same time, self-adaptive background subtraction method, the OTSU method [8], and the object segmentation method are used to subtract the background as moving target area 2. Then, the "or" algorithm is applied in these two regions to get a more perfect moving object area. Algorithm flow chart is shown in Fig. 21.3.

As shown in Fig. 21.4, part (a) is the original image of the frame from 3777 to 3779 in the sequence of video images. Figure 21.4b is the frame of 3777 of preprocessed image, the frame of 3778 is the current image, and the frame of 3779 is the result of using three-frame difference method to detect moving targets. The experimental results show that it is inaccurate to use the three-frame difference method to obtain the motion objects.

Figure 21.4c uses Kalman filtering method to extract the foreground image of the current frame 3778. This method can solve the problem of background modeling, but it can be influenced by the external light. Although the automatic OTSU threshold segmentation method reduces a portion of the noise, there are still some noise regions in the monitoring results. Thus, the inter-frame difference and the background subtraction method proposed in this chapter are combined to process Fig. 21.4c and d, with the effect shown in Fig. 21.4d. It is possible to accurately extract the complete movement of the target area.

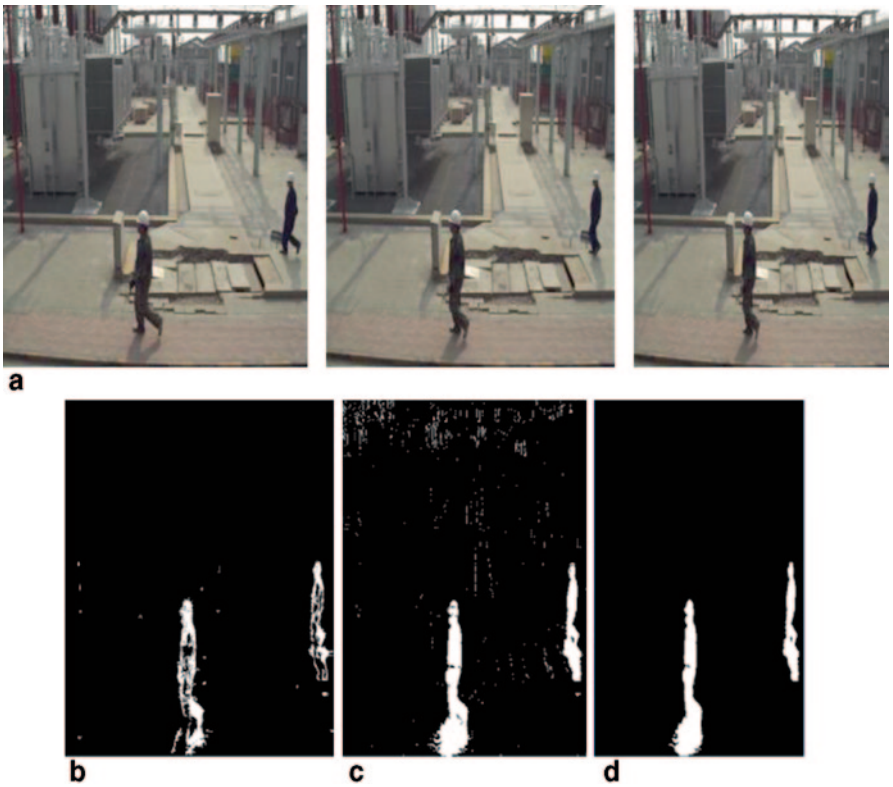


Fig. 21.4 Analysis of experimental results

21.4 A Transmission Line Intelligent Video Surveillance Method Based on the Improved Target Detection Algorithm

Based on the improved target detection algorithm proposed in this chapter, the flow chart of the intelligent security algorithm is shown in Fig. 21.5.

First, the image preprocessing based on pattern recognition includes noise reduction, image enhancement, and image segmentation. Then, it will use the improved target detection algorithm proposed in this chapter to extract the moving target foreground regions and use morphological methods and NCC-based target detection algorithm to remove noise and shadow in the moving foreground region. Finally, the type of moving target will be judged in the area of foreground region by regional unicom method.

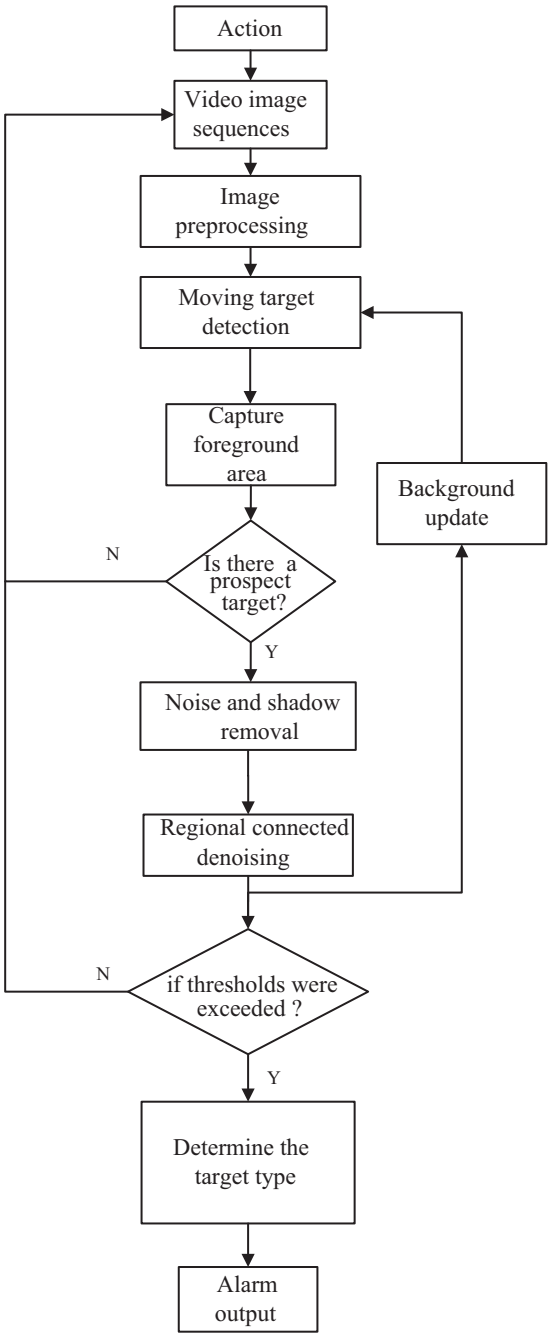


Fig. 21.5 Flow chart of transmission lines intelligent video monitoring algorithm

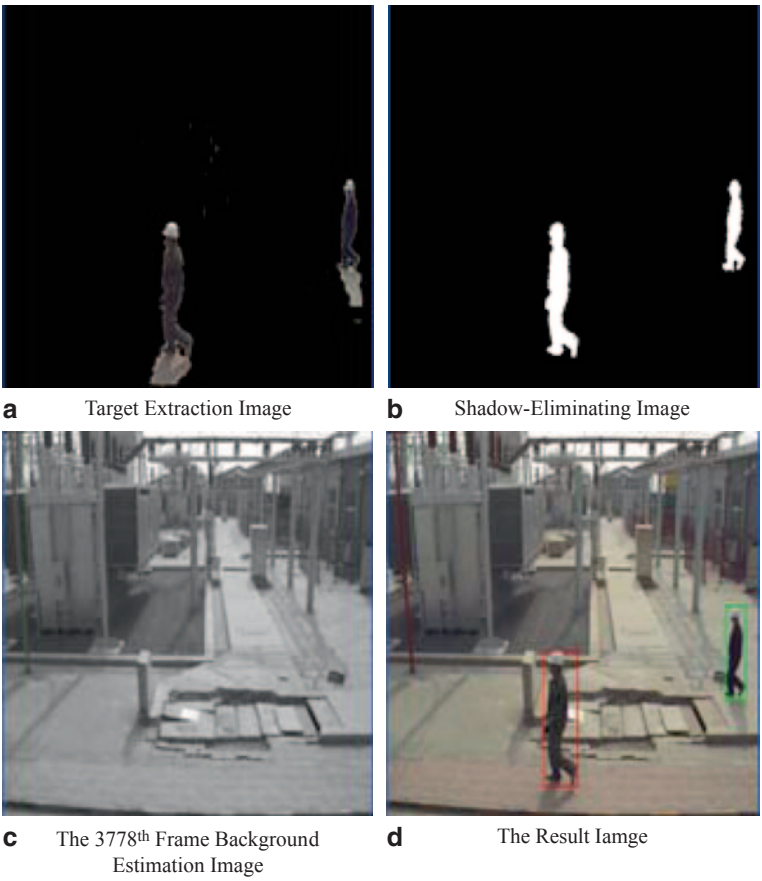


Fig. 21.6 Target detection process diagram of the 3778th frame image

Using the intelligent video surveillance algorithm improved by this chapter to process the Fig. 21.4a for the 3778-th frame image is shown in Fig. 21.6.

21.5 Case Study

The improved target detection algorithm is used to process the video images of Xinjiang YiLi Electric Power Company from a field of 220 Kv transmission line.

Use the proposed target detection algorithm to process the image. The results are represented in Figs. 21.7 and 21.8.

Fig. 21.7 This figure shows the pedestrian detection effect near a tower

Fig. 21.8 Shows the large-scale machinery and pedestrian detection effect

The experimental result shows that this method can accurately identify the pedestrians or large-scale machinery and has certain practical value.

21.6 Conclusion

The experiments demonstrate that the improved target detection algorithm by combining the three-frame difference method and the background subtraction based on Kalman proposed in this chapter can accurately detect pedestrians and large machinery. In addition, this method can observe visually the specific condition of the alarm site. It can not only make up for the previous method but also upload the alarm signal and save the cost of labor enormously. The research provides a new technical approach for the safety inspection of the transmission line.

References

1. Huang X. Transmission lines line monitoring and fault diagnosis. Beijing: China Electric Power Press; 2008. pp. 98–108. (In Chinese).
2. Hun YF. Beijing power grid transmission line damage from external situation analysis and countermeasure. North China Electr Power. 2001;1(1):12–5. (In Chinese).
3. He Y. Moving target detection and tracking algorithm based on traffic surveillance video. Zhejiang: Zhejiang University of Technology; 2011. (In Chinese).
4. Zhao J. A video image sequence moving target detection and tracking method. Beijing: Beijing Jiaotong University; 2010. (In Chinese).
5. Sun J. Background subtraction algorithm. Changsha: Changsha National University of Defense Science and Technology; 2006. (In Chinese).
6. Wu J, Du X, Liu H. A Kalman filter-based adaptive background modeling improved algorithm. Transducer Microsyst. 2012;31(1): 52–4. (In Chinese).
7. Karamann K, Brandt A. Moving object recognition using an adaptive background memory. Proc Time-Varying Image Process Mov Object ecognition. 1990;1(2):289–95.
8. Wang L, Chuan D. Otsu method in multi-threshold image segmentation. Comput Eng Des. 2008;29(11):2844–6. (In Chinese).

Chapter 22
The Exterior Type Consistency Inspection System for Smart Meter

Guiping Deng, Li Ding and Yue Guo

Abstract Based on computer vision technology, automatic control technology and laser ranging technology, the exterior type consistency inspection system for smart meters is designed and developed in this paper. The system realizes such exterior type consistency inspection items as liquid crystal display (LCD) screen, characters, bar codes, aperture size, appearance size and so on, which is designed to replace current manual detection mode. The four biaxial linkage control subsystem enables horizontal scanning, side scanning and even multi-angle scanning. The consistency inspection algorithm for exterior type based on fast template matching is used for fast and reliable character recognition and matching. The system automatically completes all the exterior type consistency detection items with fast detection speed and high accuracy and reliability.

Keywords Smart meter · Exterior type · Consistency · Inspection · Spatial orientation

22.1 Introduction

With the wide application of intelligent smart meter, its quality consistency problem has been attracting more and more attention [1]. Quality consistency is to ask the bidding sample meters and actual operating meters to be of consistent quality, including material quality, software quality and appearance type. Appearance type is mainly reflected from the size, exterior identity and LCD. The appearance type consistency is the key of smart meter quality consistency inspection. The main inspection method of smart meter appearance type consistency is visual detection by using rulers and magnifying glasses [2, 3]. With the continuous improvement in production technology, appearance type of smart meters is developing towards the direction of standardization, rationalization and high precision. The demand

G. Deng (✉) · L. Ding · Y. Guo
Metering Center, State Grid Hubei Electric Power Research Institute,
430000 Wuhan, China
e-mail: smile1248@sohu.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_22

of smart meter is dramatically increased, which puts forward the corresponding requirement of fast batch inspection.

Existing manual inspection mode has many disadvantages, such as misjudgement, low detection accuracy, low reliability, low efficiency, and inconvenience for unified computer management.

The smart meter exterior type consistency inspection is one of necessary test items for both full performance test before smart meter delivery and sampling test after smart meter batch arrival, which is specified in smart meter technical specification of State Grid of China. Therefore, intelligent recognition system of exterior type consistency based on computer technology and image processing technology has become an important issue and demand [4–6].

Based on computer vision technology, automatic control technology, laser ranging technology and image processing technology, the smart meter type consistency inspection system is designed and developed in this paper, which can effectively solve the disadvantages of current manual inspection and implement intelligent test of smart meter exterior type consistency. This system can automatically complete all kinds of exterior type consistency inspection items such as LCD screen , characters, bar codes, aperture size, appearance size and so on for both single-phase and three-phase smart meters with fast detection speed, high accuracy and high reliability. So this system will be widely used by domestic electric power companies for routine smart meter exterior type consistency inspection to replace current manual detection method.

22.2 Overall Design and Working Process of the System

Combined with machine vision and intelligent control technology, using image processing technology, optical character recognition (OCR) technology and laser ranging technology to analyze, the characteristics of smart meter appearance size, footer aperture size, nameplate identification features and LCD screen identity are gathered to form the sample library. Then, the meter to be inspected is compared with standard information in sample library, realizing the appearance consistency inspection.

The entire system consists of two parts: hardware and software. Hardware part mainly includes positioning platform with mechanical transmission part and loading table and image acquisition part, which is used to construct accurate spatial positioning environment of industrial cameras and light sources for smart meter, and to realize the image acquisition and data transmission. Figure 22.1 is the hardware framework of the system. The software part realizes the shaft drive movement control, image analysis and processing.

The system working processes are as follows:

1. Put the smart meter on loading table and keep it horizontal and fixed with positioning components. For starting the task, after successful connection of hardware with software, click on the software interface to automatically start the test work.

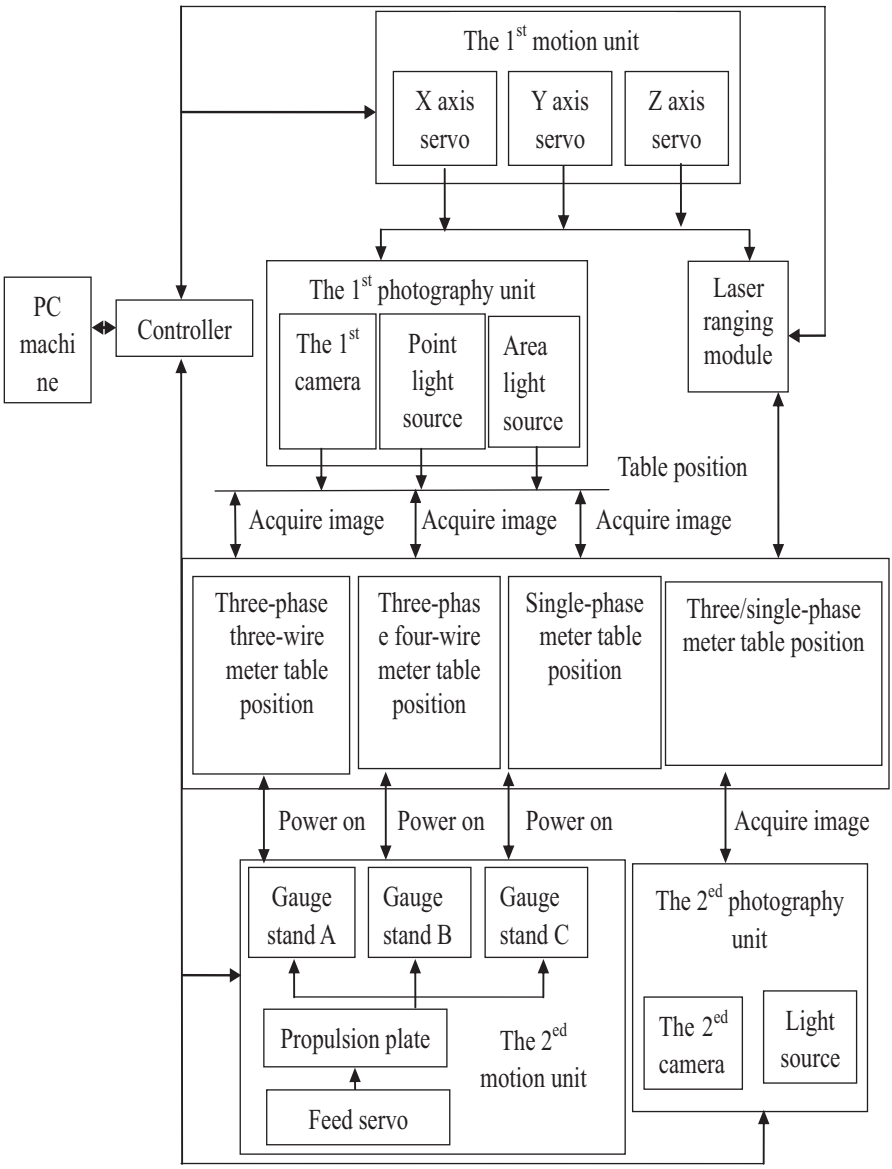


Fig. 22.1 Hardware framework of the system

2. By adjusting the four biaxial linkage control subsystem, appearance image, footer image, LCD screen information image and nameplate image of smart meter are acquired by the industrial cameras. And then the software image analysis algorithm is used to get the appearance size, footer aperture size, LCD character information features and nameplate identification character information of smart meter.

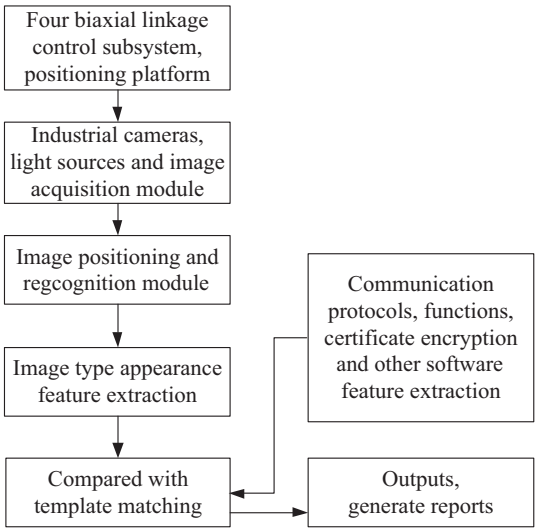


Fig. 22.2 Function modules of the system

3. Compare the exterior type information of the smart meter obtained from the field test with the information of the same specification type in the library to determine whether the appearance of the smart meter type consistency under test is qualified.

This system can realize the consistency test of appearance size, footer aperture size, LCD screenand nameplate appearance of smart meter. And it can also achieve the consistency test of software items for smart meter, such as communication protocol, function and certificate of encryption. It also includes mobile yuntai, background management, data query, history query, graphic correlation and other functions, which make it easy to use and comprehensive to test. The system is also equipped with alarm protection device. If something abnormal happens in the operating process, the alarm protection device will start automatically. The function modules of the system are shown in Fig. 22.2.

In this system, the image acquisition is based on the four biaxial linkage control subsystem. And the laser ranging method based on Steger algorithm and the character detection algorithm based on fast image template matching are developed in system inspection software. Thus, this system has high detection accuracy and reliability, as shown in Table 22.1.

22.3 Four Biaxial Linkage Control Subsystem

The four biaxial linkage control subsystem is used to control the industrial cameras, light sources, gauge stands and laser ranging module, which enables the industrial cameras to not only do horizontal scanning, but also do side and even multi-angle clear scanning. And this subsystem is one of the key parts of the hardware system. The four biaxial linkage control subsystem is designed as each biaxial with one

Table 22.1 The smart meter exterior type consistency inspection results

Test item	Standard value (mm)	Measured value (mm)	Result
Length	150.7	149.86	Qualified
Width	170	169.32	Qualified
Height	85	83.94	Qualified
LCD screen visual width	85	84.85	Qualified
LCD screen visual height	50	50.03	Qualified
Model character height	5	4.89	Qualified
Active and reactive power character height	4	3.87	Qualified
Asset barcode character height	18	18.11	Qualified
Manufacture character height	5	4.90	Qualified
Signal wiring diagram width	48	48.16	Qualified
Signal wiring diagram height	21	21.01	Qualified
Voltage and current wiring diagram width	48	48.05	Qualified
Voltage and current wiring diagram height	23.5	22.94	Qualified
Left ear screw aperture	11	10.59	Qualified
Right ear screw aperture	11	10.54	Qualified
Voltage screw aperture	4.5	4.42	Qualified
Current screw aperture	6	5.98	Qualified
Signal screw aperture	3.5	3.50	Qualified

module axis to control movement and with another guide axis to control support. This kind of design is to make the cameras and light sources move smoothly, and to avoid unclear image caused by cameras or light shaking. Driven by the industry control system, the subsystem works in accordance with the set positioning requirements. The precision of movement and control is directly related to the following image acquisition and recognition, so the precision must be accurately controlled.

The loading table is used to place smart meter under test. Due to different sizes of smart meters, the loading table plane is designed as 290×290 mm according to the maximum appearance size of smart meter. For convenient placement, the total 30 threaded holes at five rows and six columns are drilled on the plane, with the interval of 50 mm. The threaded holes are used to fix smart meters of different sizes.

22.4 Key Algorithms of System Inspection Software

22.4.1 Laser Ranging Based on Steger Algorithm

When detecting the thickness of smart meter appearance, image collections do not contain thickness information. Therefore, precise laser distance measuring technology is put forward to realize the precise measurement of the thickness of the smart meter.

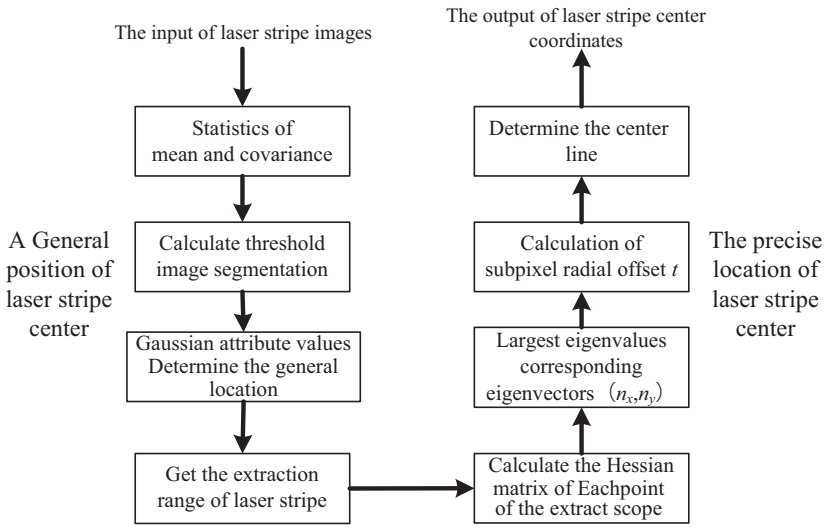


Fig. 22.3 Flow chart of laser stripe center detection

In active binocular stereo vision system based on laser stripe, the laser stripe is distorted because of the modulation of deep changes of the surface. The distortion degree contains the information of depth of the object surface under test. In order to get the depth information, the accurate position of the light stripe center must be acquired from the image containing light. As a result, in active binocular stereo vision system based on laser stripe, a very important task is to accurately extract the center of laser stripe.

In the laser stripe center detection algorithm of high precision, Steger algorithm is the best in terms of accuracy and adaptability, but the shortcoming of taking long time restricts its wide application. Therefore, this paper puts forward applying the advantages of fast detection speed of extreme value method to Steger algorithm, which greatly improve the detection speed of Steger algorithm under the premise of the same precision. Figure 22.3 is the flow chart of laser stripe center detection.

22.4.2 Character Detection Algorithm Based on Fast Image Template Matching

Current character detection algorithms are mainly divided into two categories: the reference image and the non-reference image [7]. Reference image algorithm is given the standard image of the character. The target character image area needs to be found in the whole image under test. This method is enjoying a wide application because of its easy implementation and high recognition rate. Since its shortcoming is the complexity, it is not suitable to be applied in the system which demands high detection efficiency. Thus, the OCR algorithm based on fast image template

matching is put forward, in order to realize character inspection of nameplate identification, bar codes and LCD screen of smart meter.

The basic principle of the algorithm is to use the character image in the standard library as a template, and then to search for the template image at the embedded area of the tested image. All characters of the tested image can be detected one by one, or the key area needs to be searched according to demands. Identify the word image first, and then do precise image matching. According to the similarity of the standard template and the tested image, it is determined whether there are characters of consistent type in the region, and outputs the character information.

22.5 Conclusion

Based on computer vision technology, automatic control technology, laser ranging technology and image processing technology, this paper has designed and developed the smart meter exterior type consistency inspection system. The system has high accuracy and reliability with the image acquisition based on the four biaxial linkage control subsystem, the laser ranging appearance size based on Steger algorithm and the character detection algorithm based on fast image template matching.

This system can automatically complete all kinds of smart meter exterior type consistency inspection items without manual intervention. The system has effectively solved problems such as detection efficiency, manual misjudgement and intelligent inspection. The detection items can be flexibly configured without modifying software source program, which improves its usefulness and flexibility. With large-scale smart meter used in power user electric energy data acquire system, this system will be widely used by domestic electric power companies to replace current manual detection method and has a promising prospect of application and spreading.

References

1. Zhang L. Smart meters inspection device design. Hebei: North China Electric Power University; 2012. (In Chinese).
2. Liang W. Design of smart meter visual inspection system based on machine vision. Electr Meas Instrum. 2013;50(574):64–8. (In Chinese).
3. Shen H. Bearing defect inspection based on machine vision. Measurement. 2012;45:719–33.
4. Xie G. The smart meter LCD screen appearance quality detection approach in the low-contrast environment. Comput Eng Appl. 2014;50(2):247–51. (In Chinese).
5. Bai Q. Research of automatic recognition of digital meter reading based on intelligent image processing. 2010 2nd International Conference on Computer Engineering and Technology, IEEE; 2012. p. 619–23.
6. Shao HL. Intelligent automation and control technology based on energy meter testing and storage system construction model. Appl Mech Mater. 2014;455:402–7.
7. Wang J. Research on the character detection and identification. Wuhan: Huazhong University of Science and Technology; 2007. (In Chinese).

Chapter 23
Novel Fault Isolation Method for Distributed Regional Longitudinal Protection in Smart Distribution Grid

Fanxiu Fang and Wei Cong

Abstract Considering the unique fault detection requirements of smart distribution grid, a distributed regional protection system is proposed. The distributed generations (DGs), which are widely installed in smart distribution grid, are considered to impact the power system protection performance, especially the operation of overcurrent relays that are widely used in traditional distribution grid. This chapter revises the principle of overcurrent protection that can reflect the phase-to-phase fault of next line. Then, the smart distributed terminal unit (SDTU) can use the overcurrent information and direction information to judge the fault to be at upstream or downstream. With protection-associated areas known, SDTU can exchange information with forward associated area to complete the fault isolation function. The algorithm has high selectivity and reliability through simulation verification.

Keywords New principle of overcurrent · Associated areas · Fault isolation

23.1 Introduction

Nowadays, an increasing number of renewable resources are connected to distribution grid to satisfy the customer's energy demand and decrease the use of fossil fuel [1]. However, the trend also brings challenges for traditional protection system. Traditional distribution networks are radial and unidirectional that the power flows from power source down to customers, but not vice versa. On the contrary, when DGs are connected to distribution grid, the magnitude and direction of current changed as a result of the bidirectional power flow since the DGs inject power to the grid. Under this condition, traditional protection cannot detect and isolate faults anymore; in other words, traditional fault detection, isolation methods for distribution grid will be invalid and unable to adapt to the requirements of smart distribution grid [2, 3].

F. Fang (✉) · W. Cong
Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University)
Ministry of Education, 250061 Jinan, China
e-mail: fangfanxiu@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_23

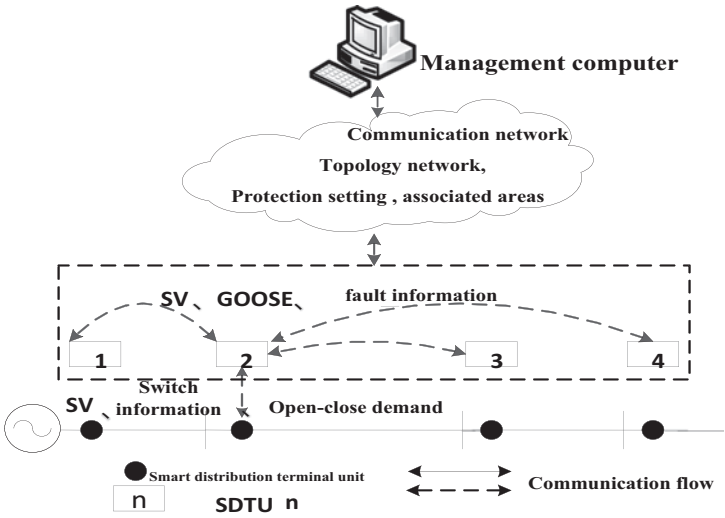


Fig. 23.1 Structure of area of longitudinal protection system

According to these problems, a distributed regional longitudinal protection system is proposed in this chapter. It uses pilot protection in distribution grid which has been widely used in transmission line. Owning to the large amount of branches and various types of DGs in distribution grid, definite associated area should be divided for each SDTU [4]. This chapter mainly presents a fault isolation method that each SDTU communicates with associated area to determine whether to break or not. The information exchange among SDTUs is fault synthetic information, which suggests the fault location. This proposed method overcomes disadvantages of current protection coordinating by time, improves reliability, and accelerates the action speed of relay protection.

23.2 Distributed Regional Longitudinal Protection System

23.2.1 Structure of Distributed Regional Longitudinal Protection System

In this chapter, in order to well illustrate the fault isolation method, the distributed regional longitudinal protection system will be briefly introduced. The protection system consists of three components: master computer, communication network, and SDTUs, as shown in Fig. 23.1.

The master computer can divide the protection zone and associated areas for each SDTU. It allocates the associated areas and sends the information to SDTUs as long as the network topology changes.

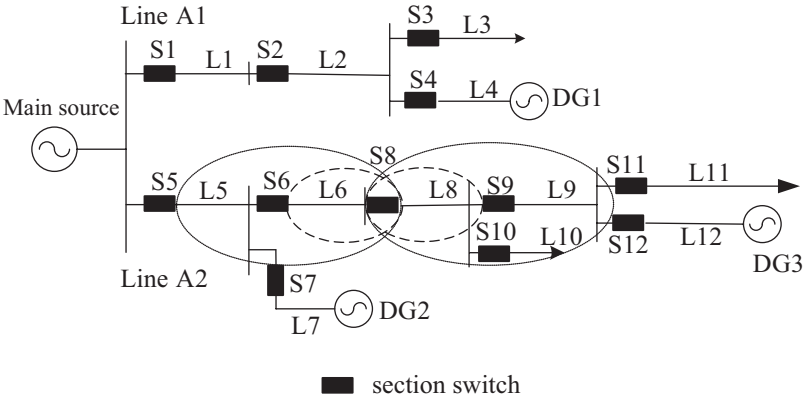


Fig. 23.2 Associated area of one SDTU

The communication network provides peer-to-peer communication and client–server communication. To be specific, a SDTU can communicate not only with other SDTUs but also with the master computer.

The primary functions of SDTU are status monitoring, fault detection, and switching control, which are realized by collecting local electrical information and communicating with other SDTUs to obtain electrical information of other nodes through peer-to-peer communication. Through embedded fault detection and fault isolation algorithm, it can make decision that whether the controlled breaker should trip or not.

23.2.2 Associated Area of SDTU

Unlike breakers in transmission line which protect only one line, each SDTU in distribution grid protects one or more areas and communicates with one or more SDTUs. It is unpractical for SDTU to communicate with every SDTU in distribution grid; thus, the protection-associated areas must be defined first.

Each terminal has its associated areas, including forward associated area and reverse associated area. The direction from main source down to DGs or loads is defined as the forward direction. One SDTU forward associated area is an SDTU which must trip to isolate the forward nearest line's fault.

Just as shown in Fig. 23.2, the SDTU9 and SDTU10 are the forward associated areas of SDTU8; the SDTU6 is the reverse associated area. Once SDTU8 cannot get the information of associated areas, it can search the backup protection-associated areas. For example, SDTU11 and SDTU12 are the forward backup protection-associated areas, while the SDTU5 is the reverse backup protection-associated area.

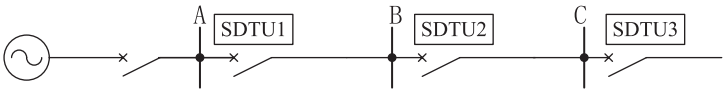


Fig. 23.3 New setting calculation of overcurrent protection

23.3 Fault Isolation

23.3.1 New Setting Calculation of Overcurrent Protection-Associated Area of SDTU

As generally accepted, the overcurrent protection is so of high sensitivity that it has large protection scope [5, 6]. Therefore, improper overcurrent setting may induce redundant communication among SDTUs, resulting in unnecessary calculations, and as a result slow down the speed of protection. In order to simplify the communication and enhance the process efficiency, a modified principle of overcurrent setting is defined.

Just as SDTU1 in Fig. 23.3, its new overcurrent setting is shown in Eq. 23.1.

$$I_{1set} = \frac{\sqrt{3}}{2} \frac{E_{\varphi}}{Z_S + Z_1 + Z_2} \frac{1}{K_{sen}} \tag{23.1}$$

Z_S is the impedance of the grid and primary transform;
Z_1 is the impedance of the line AB;
Z_2 is the impedance of the line BC;
K_{sen} is the sensitivity factor.

When the SDTU does not have lower stage line, the conventional overcurrent setting calculation method will be employed. From the new principle of overcurrent protection, it can be known that the lower stage line can be protected, meaning that its sensitivity is enough for protection.

23.3.2 Fault Synthetic Information

SDTU collects local relays' information, including direction relay and overcurrent relay. The direction relay output is defined as follows (Eq. 23.2).

$$d = \begin{cases} 1 & forward\ \ fault \\ -1 & reverse\ \ fault \\ 0 & not\ operation \end{cases} \tag{23.2}$$

The overcurrent relay output can be defined in Eq. 23.3.

$$c = \begin{cases} 1 & over\ current\ operate \\ 0 & no\ operate \end{cases} \tag{23.3}$$

Table 23.1 Fault synthetic information

Fault synthetic information (s)	(*d*, *c*)
1	(1,1), (0,1)
−1	(−1,1), (−1,0)
0	(1,0), (0,0)

Based on these results, SDTU can figure out fault synthetic information in Table 23.1.

Assume that the direction relay has higher sensitivity and lower reliability than overcurrent relay.

If fault synthetic information equals 1, it means the fault occurs in the forward line; if fault synthetic information equals −1, it means fault occurs in the reverse direction line. In addition, if it equals 0, it means this fault does not relate to it.

23.3.3 Fault Isolation

In order to isolate the fault, when the fault synthetic information of SDTU equals 1, it must communicate with its forward associated area in Eq. 23.4. Once SDTU does not have forward associated area, it can operate immediately.

$$b_i = \overline{\sum_{k=1}^{n} s_i \odot s_k} \tag{23.4}$$

$b_i = 1$ means SDTUi should trip to isolate fault and send tripping command to its forward associated area. $b_i = 0$ means SDTU need not operate. Example is used to illustrate the process of the proposed tripping calculation.

A simple distribution grid (as shown in Fig. 23.5) is simulated using PSCAD software to explain the method of fault isolation in detail. AC phase-to-phase short circuit fault is applied in line 4 (Table 23.2).

As shown in Fig. 23.4, SDTU4 is the forward associated area of SDTU2, and SDTU5, SDTU6, and SDTU7 are the forward associated areas of SDTU4.

$$b_2 = \sum_{k=1}^{1} \overline{s_2 \odot s_k} = \overline{s_2 \odot s_4} = \overline{1 \odot 1} = 0 \tag{23.5}$$

$$b_4 = \sum_{k=1}^{3} \overline{s_4 \odot s_k} = \overline{s_4 \odot s_5 + s_4 \odot s_6 + s_4 \odot s_7} = \overline{1 \odot 0 + 1 \odot 0 + 1 \odot (-1)} = 1 \tag{23.6}$$

Table 23.2 Fault synthetic information

SDTU	1	2	3	4	5	6	7	8
d	1	1	−1	1	0	0	−1	0
c	0	1	1	1	0	0	1	0
s	0	1	−1	1	0	0	−1	0

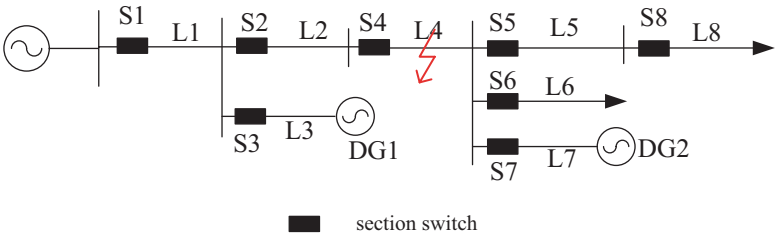


Fig. 23.4 Simple distribution network with distribution grids (DGs)

Table 23.3 New setting calculation of overcurrent protection

SDTU	800	802	806	808	812	814	850
Setting	7.938	0.9451	0.4676	0.3348	0.3341	0.3330	0.2752
SDTU	816	824	828	830	854	852	832
Setting	0.2731	0.2280	0.2257	0.1732	0.1730	0.09814	0.0444

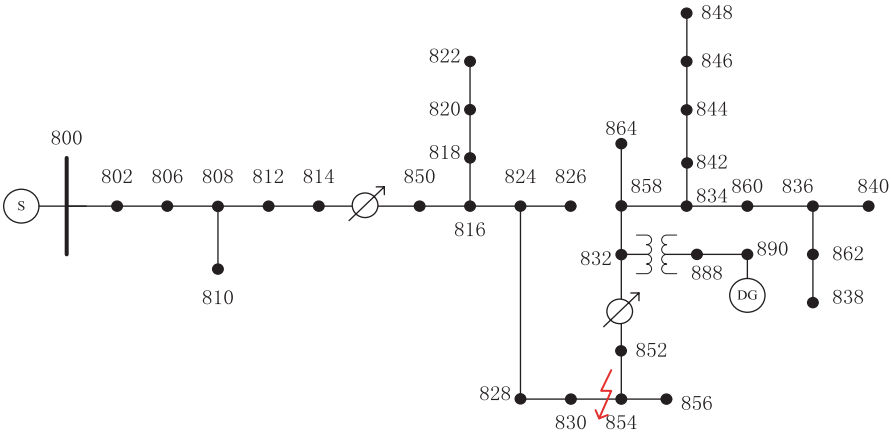


Fig. 23.5 IEEE 34-bus generator system

From Eqs. 23.5 and 23.6, it can be concluded that SDTU2 need not operate, while SDTU4 not only acts trip but also sends tripping command to SDTU5, SDTU 6, and SDTU7 to act trip.

23.4 Simulation Analysis in Distribution Grid System

In order to verify the fault isolation algorithm, the IEEE 34-bus system (as shown in Fig. 23.5) is simulated using PSCAD. AC phase-to-phase short circuit fault is set at bus 854, and as a result, the overcurrent setting of SDTUs is shown in Table 23.3.

Using the new overcurrent principle, the setting can be got as shown in Table 23.3.

Table 23.4 Output of direction, overcurrent relays, and synthetic fault information

SDTU	814	850	816	818	824	828	830	854	852	832
D	1	1	1	0	1	1	1	−1	−1	−1
C	0	1	1	0	1	1	1	0	1	1
S	0	1	1	0	1	1	1	−1	−1	−1

Table 23.5 SDTUs' associated area

SDTU	850	816	824	828	830
Associated areas	816,818	824	828	830	854

Table 23.6 Tripping command result of SDTU

SDTU	850	816	824	828	830
B	0	0	0	0	1

Simulation operating time is 3 s, the fault occurs at 1 s, and its duration time is 2 s. The output of direction relays and overcurrent relays can be acquired and shown in Table 23.4.

From Table 23.4, it can be concluded that some SDTUs' fault synthetic information equals 1. In order to detect fault, their associated areas should be got first, which are shown in Table 23.5.

According to Eq. 234, it can be decided that whether the SDTUs act trip and send message or not.

According to Table 23.6, SDTU830 should break to isolate fault. Additionally, it sends tripping command to its forward associated areas (SDTU853). The SDTU853 will break immediately after it receives the tripping command. The result is that both SDTU830 and SDTU853 act trip.

23.5 Conclusion

A distributed regional protection system is proposed in this chapter. Based on the known forward associated area, the SDTUs can exchange fault synthetic information among them to isolate faults. Through the simulation analysis, the proposed fault isolation method can isolate the fault accurately. Because fault synthetic information that combines overcurrent information and direction information is referred in protection procedure, the reliability of protection system is enhanced.

References

1. Ritter S, Rüttinger H, Bretschneider P, Westermann D. New Approaches for smart grid requirements: grid protection and optimization of distribution grid operation. Proceeding of IEEE PES General Meeting, IEEE Power and Energy Society; 2011. pp. 1–7.

2. Yin XG, Wang Y, Zhang ZZ. One-division and tripping strategy for limited wide area protection adapting to smart grid. Proc Chin Soc Electr Eng. 2010;30(7):1–7. (In Chinese).
3. Kauhaniemi K, Knmpnlained L. Impact of distributed generation on the protection of distribution networks. Proceedings of 8th IEEE international conference on developments in power system protection, IEEE Power and Energy Society; 2004. pp. 315–318.
4. Cong W, Fang FX, Shi FF. An online associated area determining method for distributed regional longitudinal protection in smart distribution grid. Autom Electr Power Syst. 2014;38(12):67–73. (In Chinese).
5. Abu-Siada A, Tin H, Masoum MAS, Qian Y. Improving the performance of smart grid overcurrent protection relays. Proceeding of Innovative Smart Grid Technologies Asia (ISGT), IEEE Power and Energy Society; 2011. pp. 1–5.
6. Gomez JC, Morcos MM. Changes in overcurrent protection paradigms of low and medium-voltage smart grids. Proceeding of North American Power Symposium (NAPS).IEEE; 2013. pp. 1–4.

Chapter 24
The Effect of Hazheng Ultra High Voltage Direct Current Commutation Failure on Wind Power

Can Yang, Haibo Zhang, Weiyong Jiang and Yanan Li

Abstract In order to study the effect of transient overvoltage caused by DC commutation failure on the wind farms, a simplified AC-DC model of the Northwest China power grid is built based on PSCAD/EMTDC. The mechanism of AC overvoltage caused by commutation failure is studied in depth. The simulations of a single-phase grounding fault at the inverter side under different conditions are performed. The factors that influence the level of overvoltage are analyzed.

Keywords Commutation failure · Overvoltage · Off-Grid wind turbines · Grid stability

24.1 Introduction

Hami is one of the planned seven wind power bases with a capacity of more than 10,000 MW in China. Commutation failure is one of the common faults in ultra high voltage direct current (UHVDC) system. The voltage of AC system may increase too high in a short time during the fault of DC system [1–3]. Once the transient voltage is higher than the voltage protection threshold of wind farms, the wind turbines will be disconnected with the grid, which will further affect the system's frequency stability; therefore, the research on the influence of Hazheng UHVDC commutation failure on the wind power in Hami is of great significance.

C. Yang (✉) · H. Zhang
College of Electrical and Electronic Engineering, North China Electric Power University, 102206 Beijing, China
e-mail: yangcan_2008@yeah.net

H. Zhang
e-mail: zhb@ncepu.edu.cn

W. Jiang
State Power Economic Research Institute, 102209 Beijing, China
e-mail: jiangweiyong@chinasperi.sgcc.com.cn

Y. Li
e-mail: liyanan@chinasperi.sgcc.com.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_24

This paper analyzes the mechanism of AC transient overvoltage caused by DC commutation failure and studies the effect of commutation failure on wind power from three aspects of flow transferring during fault, capacity of DC power and AC system operation model. All simulation results can provide a theoretical basis for the operation and dispatch of Xinjiang power grid.

24.2 Simulation Grid

24.2.1 Simulation Software

Program PSCAD/EMTDC (Power System Computer Aided Design and Electro-Magnetic Transient in DC System) is an important power system analysis software and its main function is time-domain and frequency-domain simulation. It is typically applied to calculating variation of the system parameters over time under a disturbance and is widely used for the research on the interaction of AC-DC system. PSCAD/EMTDC allows users to create custom simulation module by Fortran, C and Matlab programming languages.

24.2.2 Model of Simulation Grid

The geographic diagram of Northwest China power grid in a small basic operation mode in 2012 is shown in Fig. 24.1. It contains AC systems of Xinjiang, Gansu and Qinghai Province. Figure 24.2 shows the simulation mode of Hazheng UHVDC. It uses 2×12-pulse valve groups per pole as the circuit designing structure. In a steady state, Hazheng UHVDC is in bipolar operation mode, the transmission power is 4000 MW, the rated DC voltage is ± 800 kV and the rated DC current is 5 kA. There are 31 generators and 124 nodes in the AC system of rectifier side. The total output of wind power is 1870 MW with 1550 MW wind power connected to 330 kV node Dunhuang, 250 MW wind power connected to 220 kV node Shanbei and 70 MW wind power connected to 110 kV node Shi Sanjiang.

24.2.3 Equivalent Model of Wind Farm

There are many methods of equating wind farms [4–5], but they almost need tedious parameters reduction and the simulation accuracy is not necessarily guaranteed. The wind farm is equated to a power source in this article, as shown in Fig. 24.3. A custom module which can convert wind power to current is established by C language. C program adopts the method of single-phase average voltage [6]. Because the main program of PSCAD/EMTDC is written with Fortran language,

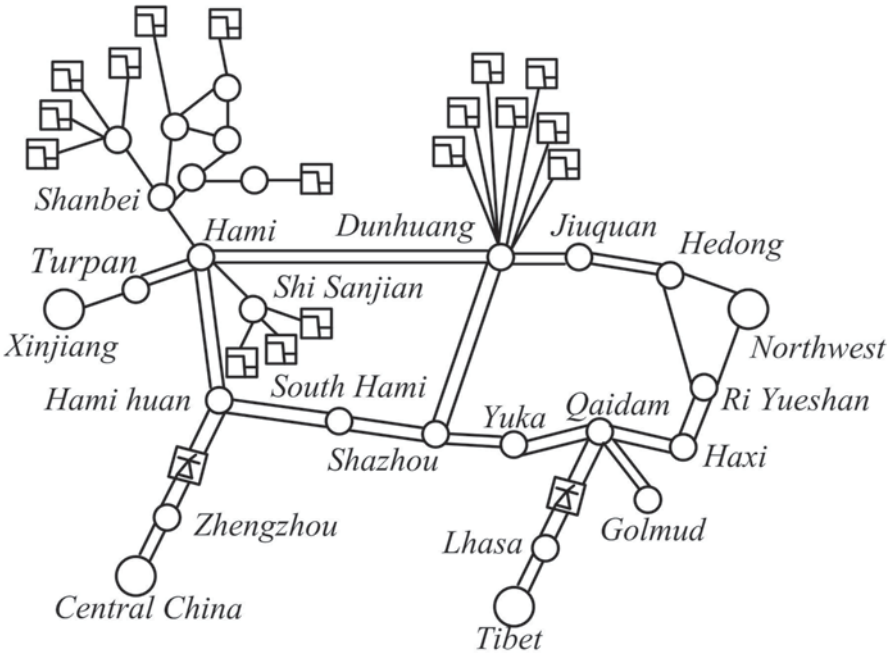
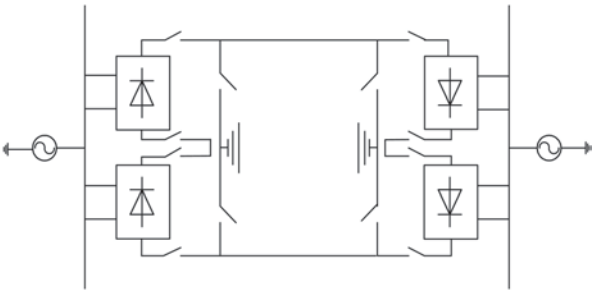


Fig. 24.1 Geographic Diagram

Fig. 24.2 Mode of Hazheng UHVDC



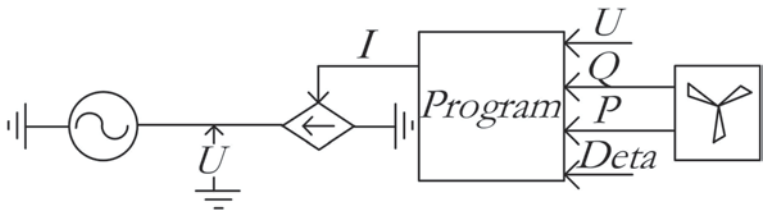


Fig. 24.3 Equivalent Mode of Wind Farm

the C program is invoked by a Fortran program in order to be embedded in the main program. The custom module, the wind turbine and the controlled current source constitute the complete dynamic wind farm model.

24.3 Simulation Research

24.3.1 Mechanism of Overvoltage Caused by Commutation Failure

This paper mainly concerns the commutation failure caused by AC system fault at the inverter side. The performance of rectifier under single-phase grounding fault of inverter bus is shown in Fig. 24.4. The fault moment is 1s and the duration is 100 ms. Thanks to the control and regulation of DC system, the power grid restores the stability eventually with no measures.

The rectifier current control system is shown in Fig. 24.5. I_{des} is the DC current setting signal from the main control layer; $I_{dO\text{–}rec}$ is the setting DC current valve, it is derived from I_{des} after Voltage Dependent Current Order Limit (VDCOL); $I_{d\text{–}rec}$ is the measured DC current; β is the advanced firing angle which can be obtained from the difference between $I_{d\text{–}rec}$ and $I_{dO\text{–}rec}$ after PI control link; according to $\beta = 180° - \alpha$, the delayed firing angle α_{rec} can be derived from 180°minus β.

After the failure is cleared at 2.1s, VDCOL increases $I_{dO\text{–}rec}$ to decrease α_{rec} so that the DC current can restore gradually; however, the speed of DC current recovering is slow while the filters in the rectifier station are providing reactive power during fault. Excess reactive power flows into AC system, which brings about an increase in AC voltage. If the inflow of reactive power is too much, the AC voltage may increase too high to trip off the wind farms and threaten the grid's stability.

24.3.2 Analysis of the Impact of Commutation Failure on Wind Power

According to the above analysis, the inflow of reactive power can easily result in AC overvoltage during the fault recovery process. The highest and stable voltage levels of the buses at wind farm terminals after DC commutation failure at 4000 MW DC power level are shown in Table 24.1. The following simulations are all assumed that a single-phase grounding fault occurs in the inverter bus and the wind turbines have the ability of high voltage ride-through (HVRT).

According to China's *Technical Rule for Connecting Wind Farm to Power System*, the high voltage protection threshold of wind farms is 1.1pu. So the wind farms connected to Shanbei and Shi Sanjiang are in the danger of off-grid. The 1550 MW wind power connected to Dunhuang can maintain the normal operation.

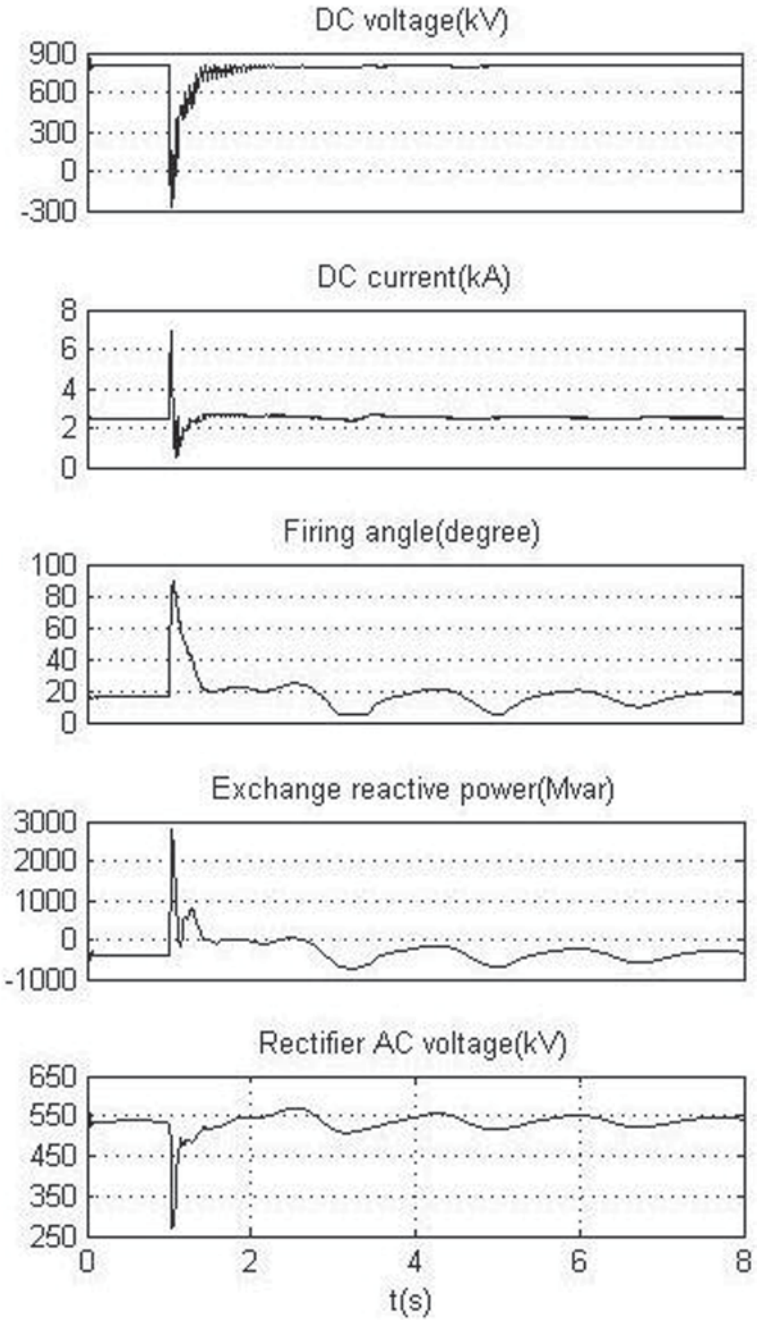


Fig. 24.4 Rectifier Simulation Results

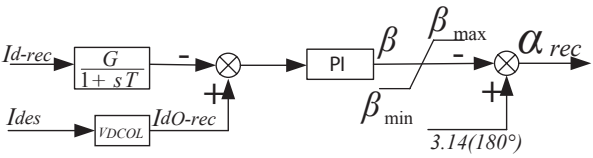


Fig. 24.5 Rectifier current control system

Table 24.1 Wind Farm Voltage Level under Commutation Failure

Buses	Highest (pu)	Stable (pu)
Dunhuang	1.0668	1.0055
Shanbei	1.1147	1.0700
Shi Sanjian	1.1235	1.0791

Table 24.2 Voltage Levels of 750 kV Nodes

AC buses	Before failure (pu)	After failure (pu)	
		Highest	Stable
Hami	0.9994	1.0400	0.9994
Dunhuang	1.0013	1.0506	1.0013
Shazhou	1.0063	1.0600	1.0063
Yuka	1.0125	1.0700	1.0125

The voltage level of the main 750 kV nodes is shown in Table 24.2. The highest transient voltage is only about 5% higher than that before failure. The results show that commutation failure does not cause a severe AC overvoltage at 4000 MW DC power level. This is because a lot of active power is transferred to Qinghai grid along with the lines Hami-Dunhuang-Shazhou-Yuka-Qiadam during the fault, which inhibits AC voltage from rising. Figure 24.6 shows the response of transmission active power of 750 kV lines to DC commutation failure. The power of Hami-Dunhuang is increased by 31% than that before fault, the power of Dunhuang-Shazhou is increased by 35%, the power of Shazhou-Yuka is increased by 49% and the power of Yuka-Chai Damu is increased by 53%.

24.3.3 Commutation Failure at Different DC Power Levels

The capacity of DC power is an important factor which affects the AC system transient voltage. Table 24.3 shows the highest AC bus transient voltage and the fluctuations of grid frequency after failure under different DC power levels.

The simulation results have shown that the more DC power is, the higher AC transient overvoltage level will be. The highest transient voltages of the three wind farms exceed 1.1pu at 6000 MW DC power level and the voltages of Shanbei and Shi Sanjian exceed 1.2pu at 8000 MW DC power level.

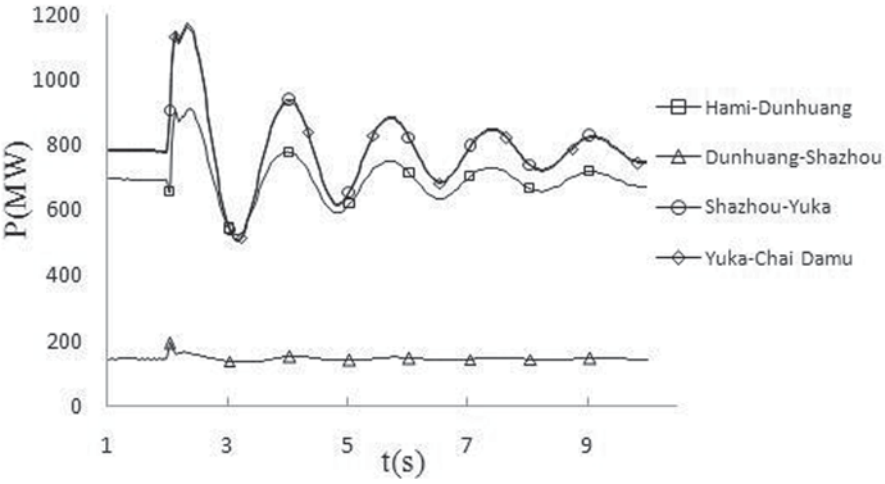


Fig. 24.6 Transmission Power of the 750 kV Lines

Table 24.3 AC System Voltage and Frequency

DC power (MW)	Voltage (pu)				
	South Hami	Hami Converter	Dunhuang	Shanbei	Shi Sanjian
4000	1.0513	1.0857	1.0551	1.1152	1.1243
6000	1.0756	1.1001	1.1472	1.1472	1.1587
8000	1.0964	1.1235	1.1849	1.2708	1.2892

24.3.4 Commutation Failure in Different AC System Operation Mode

With the expansion of the system, the transmission lines require regular overhaul in order to reduce the probability of fault. When one of the double lines is examined, all flow will be transferred to the other one, which may cause circuit overload. Table 24.4 shows the highest voltage of wind farm in different AC system N-1 mode. In the mode of the main transformer in the rectifier station N-1, Hami huan-Hami N-1 and South Hami-Hami huan N-1, both of the voltages of Shanbei and Shi Sanjian exceed 1.1pu and the voltage of Dunhuang is higher than that in other modes.

The results show that the placement of overhaul line may influence AC transient overvoltage level. It is much easier for the commutation failure to cause transient overvoltage if the maintenance line is near to the converter station. It is necessary to strengthen the security and reliability construction of the converter station and the nearby lines.

Table 24.4 Voltages of Wind Farms in Different N-1 Modes

Operation mode	Highest voltage (pu)		
	Dunhuang	Shanbei	Shi Sanjian
Transformer N-1	1.0584	1.1204	1.1286
Hami huan—Hami N-1	1.0529	1.1126	1.1213
Hami—Dunhuang N-1	0.9983	1.0704	1.0804
Dunhuang—Shazhou N-1	1.0209	1.0965	1.1052
Shazhou—South Hami N-1	1.0318	1.0813	1.0909
South Hami—Hami huan N-1	1.0558	1.1143	1.1230

24.4 Conclusion

The effect of Hazheng UHVDC commutation failure on wind power is studied based on the established AC-DC model of Northwest China power grid in this paper. The highest transient voltage levels at wind farm terminals under different conditions are checked. It is suggested enhancing the high voltage threshold of wind farms to 1.3pu so as to avoid tripping of the wind turbines. The simulation results show that the more DC power is and the nearer the maintenance line is to converter station, the higher the transient voltage will be.

References

1. Li X, Liu Y, Zhu Y, et al. Real-time simulation of dynamic performance of multi-infeed UHVDC transmission system to be connected North China Power Grid. [J] Power Syst Technol. 2005;35(8):75–80 (In Chinese)
2. Lin L, Zhang Y, Zhong Q, et al. A survey on commutation failures in multi-infeed HVDC transmission systems. [J] Power Syst Technol. 2006;30(17):40–6 (In Chinese)
3. Li X, Yi J, Li B. Simulation analysis and operation statistics of commutation failure in HVDC transmission system. [J] Power Syst Technol. 2012;36(6):266–70 (In Chinese)
4. Sun J, Jiao L, Wu J, et al. Research on multi-machine dynamic aggregation in wind farm. [J] Power Syst Technol. 2004;28(7):58–61 (In Chinese)
5. Huang M, Wan H. Simplification of wind farm model for dynamic simulation. [J] Trans China Electrotech Soc. 2009;24(9):147–52 (In Chinese)
6. Jia X, Li G, Zhao C, et al. Electromagnetic transient and electromechanical transient hybrid real-time simulation method based on RTDS/CBuilder. [J] Power Syst Technol. 2009;33(11):33–8 (In Chinese)

Chapter 25
A Maximum Power Point Tracking Method for Photovoltaic Systems

Rong Fan, XiuXia Zhang and ShunXian Bai

Abstract At present, as the use of renewable energy has attracted the people's attention, the solar power generation technology has become one of the effective ways to solve the crisis of energy and environment. In order to improve the efficiency of solar power systems, it is important to track the maximal power point (MPP). Quadratic interpolation is a method of searching optimization of a function using the sampling points function value within a certain range by means of a low interpolation polynomial approximation to the original objective function, and the polynomial optimization has been worked out and adopted to estimate the optimization target function. Based on perturbation and observation algorithm as well as the quadratic interpolation method which is introduced to narrow the scope of MPP, an improved maximum power point tracking method for PV systems is proposed. On the basis of perturbation and observation algorithm, the principle of quadratic interpolation control method is analyzed. The improved system is simulated by MATLAB. The simulation results show that the proposed method can search the maximum power point effectively and state the steady performance of the PV power systems.

Keywords PV power system · Quadratic interpolation method · Maximum power point tracking

R. Fan (✉)
School of Instrument Science and Opto-electronics Engineering, Hefei University of Technology, 230000 Hefei, China
e-mail: fanrongtg@126.com

X. Zhang · S. Bai
School of Electronics and Information Engineering, North National University, 750021 Yinchuan, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_25

25.1 Introduction

At present, as the use of renewable energy has attracted the people's attention, the solar power generation technology has become one of the effective ways to solve the crisis of energy. In order to improve the efficiency of solar power systems, it is important to track the maximal power point (MPP).

Ordinarily, the method of achieving maximum power point tracking (MPPT) control is based on the combination of hardware and software. Recently, some new control algorithms have been realized [1–3], such as the constant voltage method, the linear approximation method, the power disturbance observation method, the total conductance method, the fuzzy control method and some other MPPT control methods. In order to improve the efficiency of the solar power systems, it is important to track the MPP. The perturbation and observation control method of MPPT features few measured parameters and simpler structure. The method of MPPT features less measurement parameters, simple structure, and constant disturbance to the track maximum power point [4] while the disadvantages highlight the tracking precision and speed closely related to the tracing step length and the initial value, and moreover likely to have shock phenomenon near the maximum power point, which thus leads to miscarriage of justice. The quadratic interpolation is a method of searching optimization of a function by using the sampling points function value within a certain range by a low interpolation polynomial approximation to the original objective function. The polynomial optimization is worked out and used to estimate the optimization target function [5].

Based on the traditional MPPT control method, the perturbation and observation algorithm and the quadratic interpolation method which are introduced to narrow the scope of MPP, an improved maximum power point tracking method for PV systems is proposed in this paper.

25.1.1 Quadratic Interpolation Method

Quadratic interpolation is a method of searching optimization of function $f(x)$ by using the sampling points function value within a certain range [6] by using a low interpolation polynomial approximation to the original objective function, and the polynomial optimization is worked out and adopted to estimate the optimization target function $f(x)$.

In order to satisfy the given precision, the fitting process can be repeated. There are three points $(x_0, f(x_0))$, $(x_1, f(x_1))$, $(x_2, f(x_2))$, $(a \le x_0 < x_1 < x_2 \le b)$ in the searching rang of target function $f(x)$, which is shown in Fig. 25.1. On the condition of satisfying the nature of the parabola of $f(x)$, three points should be $f(x)_0 < f(x)_1$, $f(x)_1 > f(x)_2$. The quadratic interpolation basis function of x_0, x_1, x_2 is $l_0(x) = \dfrac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}$, $l_1(x) = \dfrac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}$, and $l_2(x) = \dfrac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$; the quadratic interpolation polynomial is [5]:

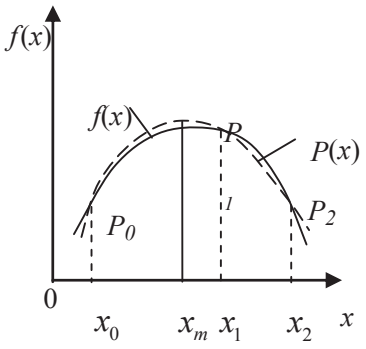


Fig. 25.1 Fitting curve of quadratic interpolation method

$$L_2(x) = l_0(x)f(x_0) + l_1(x)f(x_1) + l_2(x)f(x_2) \tag{25.1}$$

The maximum point of fitting curve is $\left(-\frac{b}{2a}, \frac{4ac-b^2}{4a}\right)$

25.1.2 Perturbation and Observation Control Method of the PV Systems

Photovoltaic cell could produce light voltage and photo-generated current if the load is joined in the external circuit in the sunlight shine, but the light voltage is neither constant nor the photo-generated current. They are both influenced by sunshine intensity and environment temperature. The relations between light voltage and photo-generated current is given by

$$I = I_{ph} - I_D - I_{sh} = I_{ph} - I_0[e^{\frac{q(V+IR_s)}{AKT}} - 1] - \frac{V + IR_s}{R_{sh}} \tag{25.2}$$

In the equation, I is photovoltaic cell output current, I_{ph} is photo-generated current, I_D is current of diode, I_{sh} is leakage current of shunt resistance R_{sh}, I_0 is reverse saturation current, q is charge constant, V is output voltage, A is p-n junction coefficient of diode, K is Boltzmann constant, and T is Kelvin temperature. According to the electronics theory [7], the physical equivalent circuit model for photovoltaic cell is shown in Fig. 25.2. The relationship between solar cell output voltage, current and power is shown in Fig. 25.3, the I-V characteristic curve and the P-V characteristic curve. The maximum output power is $P_m = I_m V_m$; thus, the MPP has characteristics of sunlight intensity and temperature variation. A traditional MPPT control method (Perturbation and Observation) is often used with the basic principle shown as below: First, an output voltage signal U_{PV} and a perturbation step length ΔU are initialized, then the voltage signal and its power value are

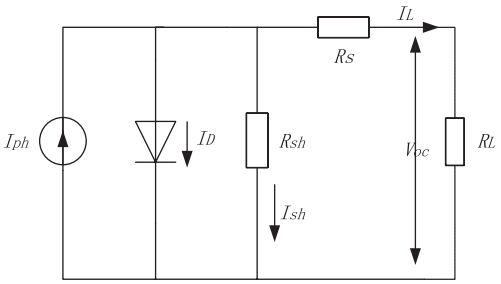


Fig. 25.2 Equivalent circuit of solar cells

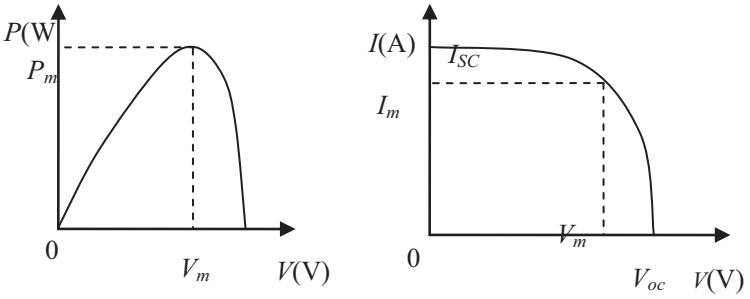


Fig. 25.3 P-V and I-V characteristic curve of solar energy battery

measured. If the relationship between the voltage and the power variation is positive, it is proved the direction is right; thus, it continues to disturbance to the same direction $(+\Delta U)$; on the contrary, if the relationship is negative, the disturbance to the adverse direction $(-\Delta U)$ [8].

25.2 Improved Maximum Power Control Algorithm of PV Systems

The perturbation and observation control method of MPPT features few measured parameters and simpler structure. The advantage perturbation and observation control method of MPPT highlights less measurement parameters, simple structure and with constant disturbance to track maximum power point. The disadvantage is the tracking precision and speed closely related to the tracing step length ΔU and the initial value U_{PV}, and moreover likely to have shock phenomenon near the maximum power point. Thus, it leads to miscarriage of justice. With the advantages of quadratic interpolation, the curve of solar power is fitted by it. Three points $(U_1, P_1), (U_2, P_2), (U_3, P_3)$ within a limited range are selected and they should meet the conditions as below:

$$(U_1 > U_2 > U_3) \,\&\, (P_2 > P_1, P_2 > P_3) \tag{25.3}$$

$$(U_1 < U_2 < U_3) \& (P_2 < P_1, P_2 < P_3) \tag{25.4}$$

Then calculate voltage of maximum power point voltage as $U_m = -\frac{b}{2a}$.

The flow chart of the improved method for MPPT is shown in Fig. 25.4:

1. Set the step of ΔU and the initial value of U, δ; the perturbation and observation algorithm is used to reach three points $(U_I, P_1), (U_2, P_2), (U_3, P_3)$; then search the MPP. Stop the search when the variation of voltage is less than δ;
2. According to the MPP as searched in step 1, select three points nearby it, and content Eq. (25.3) or Eq. (25.4), the maximum power point voltage $U_m = -\frac{b}{2a}$ is calculated;

25.2.1 Simulation Results

In this paper, a MATLAB simulation is built based on the maximum power tracking control of the small solar power system as the research object. Set $\Delta U = 0.1v$, $I_{sc} = 3.2A$, $V_{oc} = 22\text{V}$, $I_m = 2.94A$ and $V_m = 17V$ with the light intensity changing from 1000 to 500 W/m^2 at 0.15 s. The simulation diagram of PV output power based on the disturbance observation algorithms is shown in Fig. 25.5. It can be seen that power value fluctuation is larger at the beginning and $0.15s$. The results show that the proposed strategies can narrow the scope of MPP and find MPP accurately when the environment is changed, avoid the oscillation of perturbationobservation algorithms and improve the power efficiency.

The simulation diagram of PV output power based on the improved algorithms is shown in Fig. 25.6. By comparing Fig. 25.5 and Fig. 25.6, it could be seen that power value fluctuation is smaller at the beginning and 0.15 s. Both the power error and the searching time based on two control algorithms are analyzed. The results are listed in Table 25.1, which shows the PV output power based on improved control algorithm. The PV output power is 199.73 and 95.09 W, respectively, which is closer to the theoretical value than that of perturbation and observation algorithm when light intensity is changed. Table 25.1 also shows that the time to search the maximum power point. The improved control algorithm takes 0.02382 s and 0.02422 s, respectively, with less time in searching MPP when the light intensity is changed. Experimental results verify that the proposed algorithm is superior to the perturbation and observation algorithm in terms of precision and searching time.

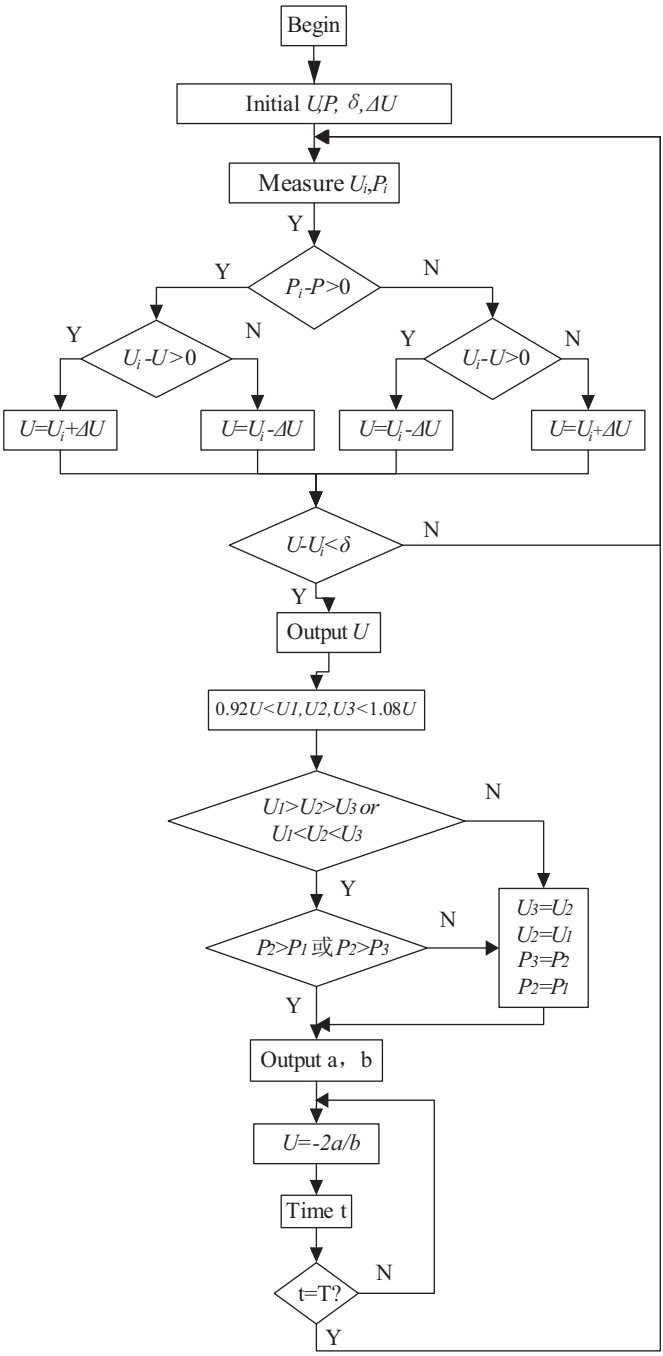


Fig. 25.4 The flow chart of improved method for MPPT

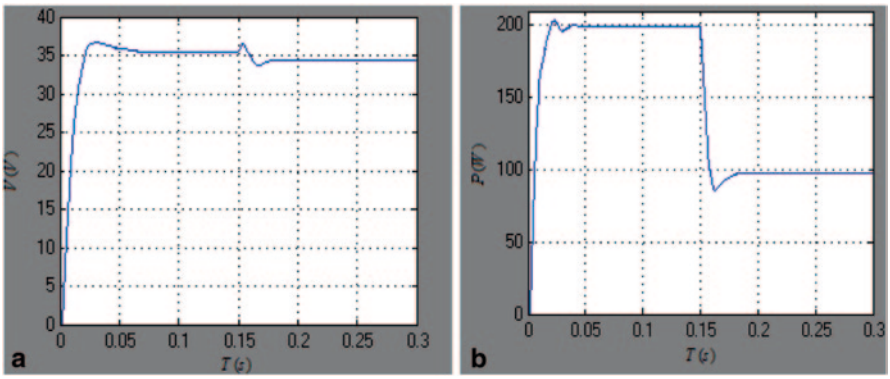


Fig. 25.5 Simulation of PV output voltage and power with light intensity change

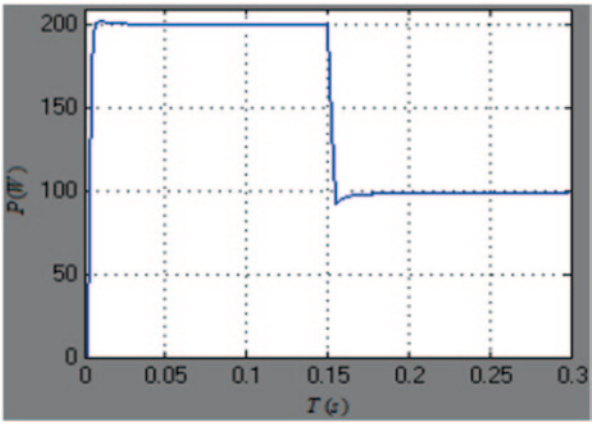


Fig. 25.6 Simulation of PV output power with light intensity change based on improved algorithm

Table 25.1 PV output power and searching time based on two algorithms

	PV output power (W)				Searching time (S)	
	1000 W/m^2	Error	500 W/m^2	Error	1000 W/m^2	500 W/m^2
P & O algorithm	195.96	2.26%	93.32	1.93%	0.05153	0.04164
Improved algorithm	199.73	0.38%	95.09	0.07%	0.02382	0.02422
Theoretical value	200.49	–	95.16	–	–	–
Time saving (S)	–	–	–	–	0.02771	0.01742

25.3 Conclusion

From the comparison, it could be drawn that the improved control algorithm could find maximum power point more quickly when the external environment condition is varied. And the shock phenomenon can be avoided with the power conversion efficiency improved at meanwhile. Experimental results verify that the improved control algorithm is superior to the perturbation and observation algorithm, and the proposed method is stable and feasible.

Acknowledgements The project was supported by the National Scientific Fund of China (No.51365001) and North National University Scientific Fund (No.2011Y029).

References

1. Hu J, Zhang J, Wu H. A novel MPPT control algorithm based on numerical calculation for PV generation systems power electronics and motion control conference, 2009. IPEMC'09. IEEE 6th International 2009. p. 2103–07.
2. Goudar MD, Patil BP, Kumar V, A review of improved maximum peak power tracking algorithms for photovoltaic systems. Int J Electr Eng Technol. 2010;1(01):2–94.
3. Qiu P, Ge B, Bi D. MPPT control for PV power generation system based on P & O algorithms and quadratic interpolation. Power Syst Prot Control. 2011;39(4):62–7.
4. Zhang L, Zhang X, Liu T, Wang E, Yang X, Wei S. Reviewed of maximum power tracking for wind-solar hybrid generation system. Electr Des Eng. 2012;20(23):29–31. (In Chinese)
5. Ou Y, Nie Y, Che M. Numerical analysis. Beijing: Higher Education Press; 2009. p. 40–67. (In Chinese)
6. Li Q. Numerical analysis basic tutorial. Beijing: Higher Education Press; 2001. p. 31–45. (In Chinese)
7. Su J, Yu S, Zhao W, Wu M, Shen Y, He H. Investigation on engineering analytical model of silicon solar cells. Act Energ Sol Sin. 2001;22(4):409–12. (In Chinese)
8. Chen Y, JIN Y. Reviewer of control technologies for wind-solar hybrid generation system. Electr Driv. 2012;42(1):3–9. (In Chinese)

Chapter 26
Motion State and Transient Analysis of the Hybrid Power Coupling Mechanism Based on Matlab and Solidworks

Sulan Han, Shufang Feng and Junlong Zhao

Abstract The hybrid vehicle is an important development of the automobile industry because of its capabilities in reducing pollution and saving energy. The key component of its power system is the power coupling mechanism which determines the power efficiency. Therefore, the planetary gear power coupling mechanism which is applied in heavy hybrid vehicles is studied. By using Matlab, an entire vehicle simulation model is established to obtain the output rotating speed and torques under different working conditions. And with the application of Solidworks, a simulation model of the coupling mechanism is set up to obtain its dangerous working period. Then, the simulation data are used to carry out the transient analysis of the coupling mechanism working during the dangerous period. The research results indicate that this method is simple and reliable and can provide technical support for strength-optimized design of the mechanism.

Keywords Matlab · Solidworks · Hybrid · Coupling mechanism · Transient analysis

26.1 Introduction

Transient dynamic analysis is a method to determine the dynamic response of structure under arbitrary time-varying loads. It can be used to analyze the time-varying stress, strain, and displacement of structure under the arbitrary combination of steady loads, transient loads, and harmonic loads.

Hybrid vehicles generally refer to cars which have an assembled power system with internal combustion engine, electric motor, and storage battery. According to the drive coupling mode, the hybrid power system can be divided into parallel hybrid, series hybrid, and mixed hybrid. Due to the flexibility in adjusting the output

S. Han (✉) · J. Zhao
North China University of Water Resources and Electric Power, 450045 Zhengzhou, China
e-mail: hansulan@ncwu.edu.cn

S. Feng
Xinxiang Xinovo Machinery Co., Ltd, 453000 Xinxiang, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_26

power of combustion engine and electric motor under various circumstances, the mixed hybrid power system has been widely applied. There are two typical mixed hybrid system layout schemes: planetary gear and clutch. The major part of the planetary gear mixed hybrid vehicle is its coupling mechanism [1].

The planetary gear hybrid coupling mechanism used in heavy hybrid vehicles is studied, and, the transient analysis is applied to analyze its working states. First, Matlab is used to establish an entire vehicle simulation model to get the output rotating speeds and torques under different working conditions, and the dangerous part of the coupling mechanism can be obtained through theoretical calculation. Second, the three-dimensional modeling, virtual assembling, and kinematics simulation of the hybrid coupling mechanism are executed by using Solidworks to get the dangerous working period. Finally, the simulation data are used to carry out the transient analysis of the coupling mechanism working during its dangerous period. Therefore, the achievements will provide technical support for strength-optimized design of the mechanism.

26.2 Mixed Hybrid Vehicle Coupling Mechanism

The working principle of a hybrid vehicle is to couple the power from the engine and motor and export it to the driving wheels. The major part of a mixed hybrid vehicle is its coupling mechanism.

Because brake and clutch are merely the assisting parts to switch the motion mode, the core part of the coupling mechanism is the planetary line. Due to the complication of the actual model which is hardly used to analyze the strain and stress of the whole coupling mechanism, a simplified model has been developed. With neglecting the rotating friction between them and the planet gears, removing bearings, hubs, and unifying tooth width, the simplified model is composed by sun gear, planet carrier, ring gear, and planetary gear, considering the planetary gear shafts and the planetary disk as a whole.

When driving a vehicle, the power from the engine is divided into two parts: one part is used to drive the vehicle, and the other is used to drive the motor to generate electric power which will be stored in the storage battery. Thus, the earning power of the engine is used. During normal driving, the driving power is provided by the motor from the electric power stored in the storage battery. While driving at lower speed and with higher driving torque (i.e., while climbing), the driving torque provided by the motor is inadequate, and the vehicle will enter into the hybrid driving mode; the engine will start to provide power to drive the vehicle at the same time through the coupling mechanism. It will not only ensure the higher power requirements of the vehicle, but also improve fuel economy and avoidwaste of energy [2].

26.3 The Simulation of the Coupling Mechanism Working Modes

26.3.1 Establishing the Vehicle Model Based on Matlab

The SimDriveline of Matlab is applied to select battery, generator, motor, and internal combustion engine modules to build the entire vehicle frame based on the actual structure. Using the electric system provided by the components library of SimDriveline, all modules are connected together to establish the entire vehicle simulation model, as shown in Fig. 26.1.

26.3.2 Analysis of the Vehicle when Accelerating to 90 km/h

Using the simulation model established in Matlab, the different working conditions of vehicle can be simulated. By simulating the working condition of accelerating to 90 km/h, the experiment result is achieved. Based on the specific time-varying rotating speed of the sun gear and planet carrier, the torque of the sun gear and the experimental data are cataloged for import into Excel to draw the charts. The varying charts of the vehicle speed, the rotating speed of the sun gear and planet carrier, and the torque of the sun gear are shown in Figs. 26.2, 26.3, and 26.4, respectively.

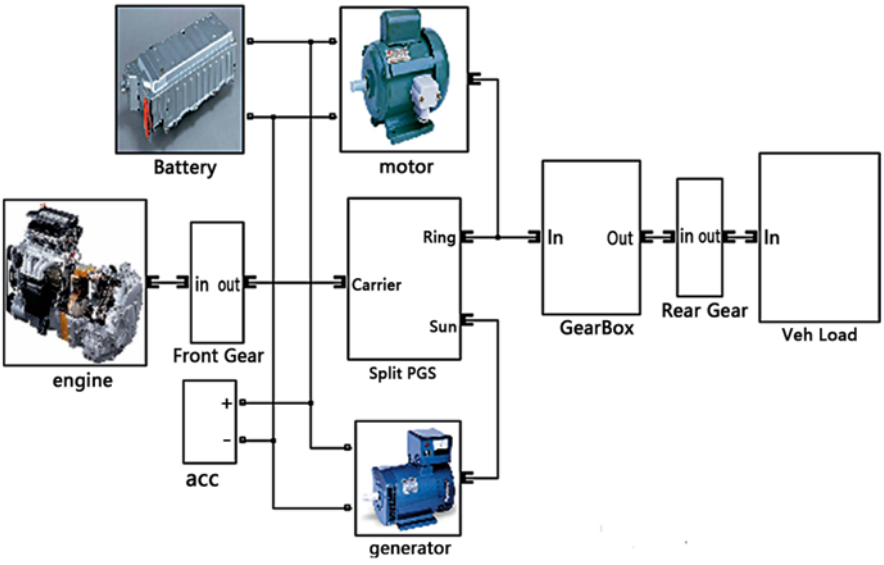


Fig 26.1 Vehicle simulation model

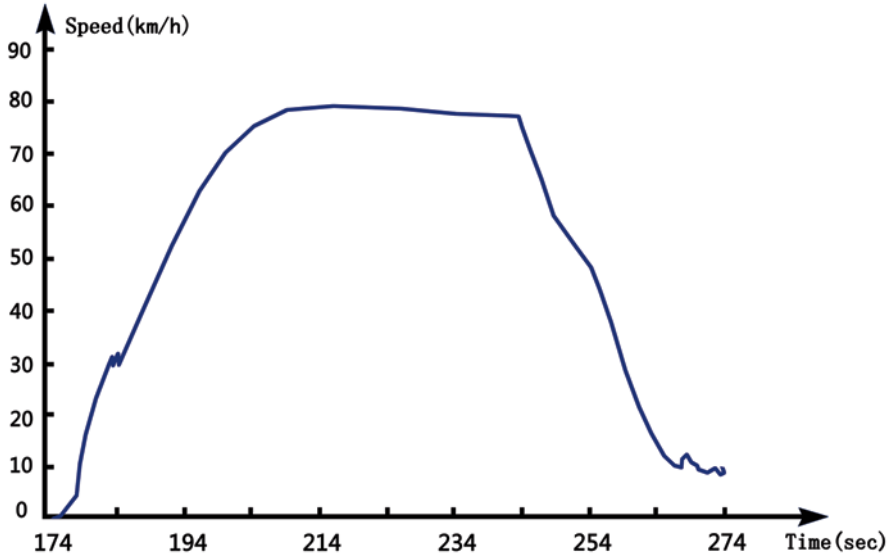


Fig. 26.2 The changing vehicle speed when accelerating to 90 km/h

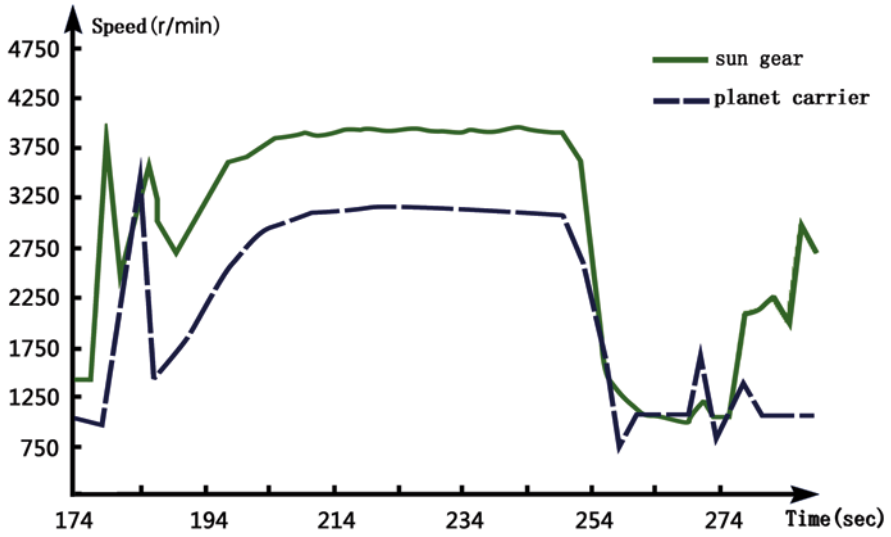


Fig. 26.3 The rotating speed of sun gear and planet carrier when accelerating to 90Km/h

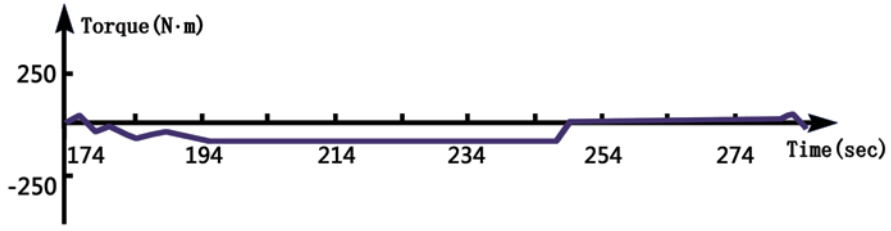


Fig. 26.4 The torque of sun gear when accelerating to 90Km/h

Because of the large number of data, we represent only the typical data within the range of 174–274 s, and the dynamic characteristics of the coupling mechanism will be studied in this range. Set the start time to 0 and save it as .txt file for the future application. From the figures given below, we can see that the three parameters vary at different times, especially the rotating speed of sun gear, which is related to the input of planet carrier.

26.4 Theoretical Calculation of the Coupling Mechanism

The research object is a single-row planetary gear, the modulus m = 3, the number of sun gear teeth $Z_1 = 25$, the number of ring gear teeth $Z_2 = 63$, and the torque is 900 N·M, loading on the planet carrier.

The gear surface contact stress calculation formula is [3]:

$$\sigma_H = Z_H Z_E \sqrt{\frac{KF_t(\mu \pm 1)}{db\mu}} \tag{26.1}$$

The tooth root bending stress calculation formula is [3]:

$$\sigma_F = \frac{KF_t}{bm} \times Y_{Fa} \times Y_{Sa} \tag{26.2}$$

In formula (1), Z_H is the node region coefficient, Z_E is the elastic coefficient, K is the load coefficient $K = KAKVK\beta$ (where K_A is the usage coefficient, K_V is the dynamic load factor, and K_β is the tooth load distribution factor), b is the tooth width, $\mu = Z_2/Z_1$, and the positive sign is for external mesh, negative sign is for internal mesh.

In formula (2), Y_{Fa} is the tooth shape coefficient, Y_{Sa} is the stress correction factor, and K is the load coefficient. The value and meaning equals to the K of the contact stress, and in static analysis, $K = 1$.

Calculate the stress of sun gear and ring gear according to the above analysis, as shown in Table 26.1.

The calculation shows that the maximum surface contact stress appears at the ring gear, and the bending stress of the sun gear is similar to that of the ring gear. Considering that the coupling mechanism of the planetary row is a mechanism with an increasing speed and a decreasing torsion [3], the torque and stress of the ring gear should be larger than those of the remaining components; thus, only the "dangerous" gear ring needs transient analysis and check.

Table 26.1 The stress of sun gear and ring gear

Stress	Sun gear (N·M)	Ring gear(N·M)
Contact stress	256	275
Bending stress	315	314

26.5 3D Modeling, Virtual Assembly and Transient Analysis of Coupling Mechanism

26.5.1 3D Modeling and Virtual Assembly

Create a new part of drawing in Solidworks and save it as “planetary gear”. If z (the number of gear teeth) and m (the number of modulus) are already given, the 3D model of the planet gear, sun gear, planet carrier, and planetary axle can be created successively [4].

Create a new assembly file and save it as “coupling mechanism.” At the beginning of assembling, the sun gear will be inserted first and assembled correspondingly due to the matching relation; then, the carrier, planet wheels, and ring gear will be inserted successively to establish the coupling mechanism, as shown in Fig. 26.5 [5].

26.6 Transient Analysis of Coupling Mechanism when Accelerating to 90 km/h

Use the rotating speed and torque achieved in section 26.3.2 as the input data to make transient analysis of the above 3D model, and the processes are as follows:

1. Open Solidworks and select Solidworks Motion and Solidworks Simulation in plug-ins.
2. Open the coupling mechanism assembly, execute "Motion Studies" and select the type of study as “Motion Analysis.”

Fig. 26.5 The assembling drawing of the coupling mechanism

3. Contact setting: Since the organization is more complex, it requires the definition of 12 contacts.
4. Torque and speed definition: There are three data need to be defined in each working conditions; they are, respectively, the speed, torque of sun wheel, and the speed of the planet carrier.
5. Simulation setting: In order to get the time-varying stress–strain curve, add simulation settings to sun gear, planetary gear, and the ring gear.
6. Calculation.

Because of the large number of data points, merely the torque of the planet carrier within the time interval with maximum value is chosen to calculate the simulation results. Execute the "Results and Plots" command; the diagrams of the corresponding torque, the speed of gear ring, and the sun gear torque will be generated after a certain time of calculating. In the gear ring torque diagram, we will find that the larger carrier torque interval is within the range of 178–183 s, as shown in Fig. 26.6.

Furthermore, take the rotation speed and torque of the sun gear, as well as the planet carrier speed as input data, and execute the "computer simulation results" command; the stress nephogram of the coupling mechanism under two different working conditions will be obtained, as shown in Fig. 26.7.

The above study indicates that the theoretical values are consistent with the simulation results and in accord with the expected material properties. The stress error is in the allowable range.

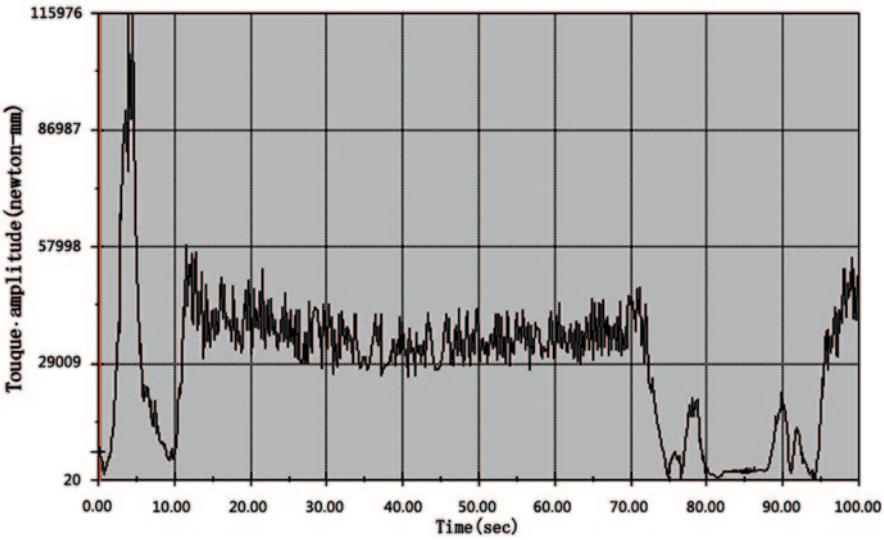


Fig. 26.6 Carrier torque when accelerating to 90 km/h

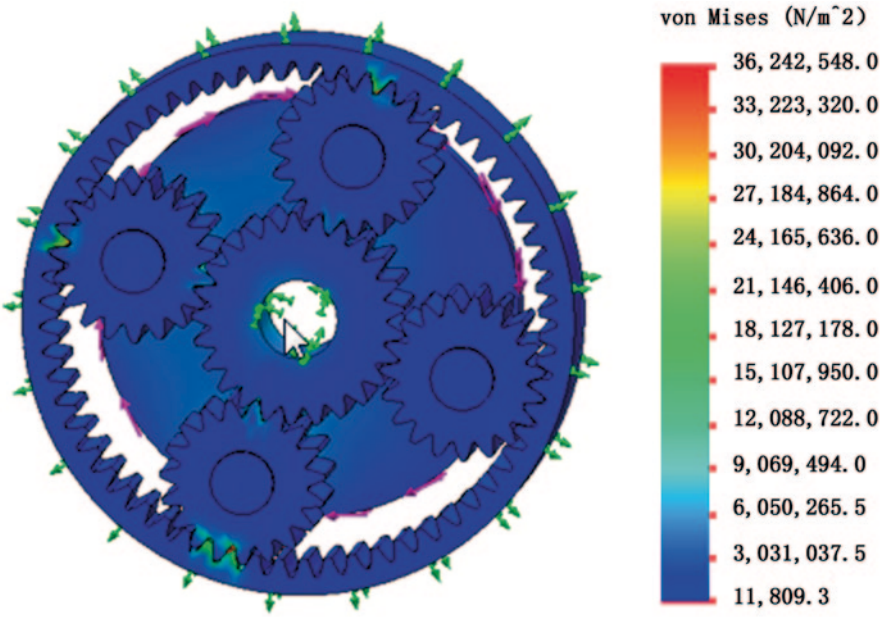


Fig. 26.7 Stress nephogram of coupling mechanism under two working conditions

26.7 Conclusion

Matlab and Solidworks are jointly used in this chapter, and Matlab is used to simulate the output forces. Based on the theoretical calculation for the planetary gear coupling mechanism, the Motion module is applied to carry out the transient analysis for the stress and strain of the "dangerous" part, the ring gear. By using transient analysis, this method has overcome the limitation of static analysis in verifying the strength of designed structure. And the smaller dangerous range can efficiently reduce the dynamic data input. The research results indicate its simplicity and reliability.

References

1. Wu G, Qin D, et al. Research on power transmission systematic scheme and matching of parameters of the hybrid power planetary gear train. J Mach Des. 2009;26(06):60–3. (In Chinese)
2. Chen L. Analysis of hybrid electric vehicle power and economy. Wuhan: Wuhan University of Technology; 2008. p. 10–2. (In Chinese)
3. Pu L, Ji M. Design of machinery. 8th ed.Beijing: Higher Education Press; 2010. p. 201–3. (In Chinese)
4. Han S, Yuan L. The improvement of direct modeling method to generate gear. J North China Water Resour ang Electr Power. 2013;34(06):100–1. (In Chinese)
5. Cui L, Gong X, et al. Movement simulation of planetary gear transmission based on Solidworks. J Mech Transm. 2008;32(05):59–61. (In Chinese)

Chapter 27
Robust Variance Control Algorithm for the Dish Solar Generation Tracking System

Chen Chen, Pingkang Li and Xiandeng Ping

Abstract The problem of strong winds with sands is encountered in the application environment of the dish solar thermal power technology. This chapter presents the modeling robust variance control (RVC) of the dish solar generation tracker. The structure and the operation principle of the dish solar tracking system are discussed. On this basis, the state space model of tracking servo system with the random disturbances of wind load and system parameter uncertainties is built. In order to overcome the disturbances and uncertainty, the robust variance constraints algorithm is applied to the tracking system in the form of the state feedback matrix. This chapter also analyzes the influence of circle centre location which is the key design parameter of the robust variance control on the output angle variance. Compared with the existing results, the simulation results show that the robust variance control has better robustness and higher control accuracy.

Keywords Tracking system · Dish solar thermal power generating · Robust variance control · Wind disturbances · System uncertainties

27.1 Introduction

With the rapid development of the new energy technology, the solar thermal power generation has become an important part of distributed generation. According to different ways of heat accumulation, the solar thermal power generation is divided into tower type, dish type, trough type, and Fresnel type. Among them, the dish type has the highest photoelectric conversion efficiency [1]. The key factor of the high photoelectric conversion efficiency of dish solar system is how to focus the light reflection on the central region of the receiver steadily. Furthermore, the tracking system mainly determines the position of the focus on the receiver. Therefore, the control performance of the tracking system affects the thermal efficiency of the receiver. Dish solar systems can be used as a separate power supply in distributed

P. Li (✉) · C. Chen · X. Ping
School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, 100044 Beijing, China
e-mail: pkli@bjtu.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_27

generation or as a subsystem connected to the grid. Most dish solar systems have to be located in the area with rich solar energy resources [2]. The characteristics of strong winds with sands, large temperature difference, and severe pollution require the tracking system to resist wind, sand, and dust. Therefore, the study of high precision with robust performance dish solar tracking control technology is of great significance.

Recently the control strategies of the solar tracking system have been discussed in many works. Fuzzy, sliding, phase lead, and other control strategies are brought to analyze the tracking control system to improve tracking accuracy [3–5]; however, the influence of the disturbance from wind loads on the thermal efficiency has not been analyzed. The disturbance causes the pointing error to reduce the tracking performance. Thus, it is very necessary for the dish solar tracking system to design the controller with high precision, good stability, and strong anti-disturbance to overcome the external wind load disturbance.

In this chapter, the operation principle and composition of the dish solar tracking control system are introduced in Sect. 2. Considering the uncertainties and the external random disturbance of the tracking system, the state space model of the tracking servo system is established. In Sect. 3, the state feedback controller based robust variance control (RVC) algorithm is present. In Sect. 4, the affecting relation between the circle centre location and the output angle variance is analyzed. Finally, the conclusion is presented in Sect. 5.

27.2 Dish Solar Tracking Control System

The solar thermal power generation is composed of parabolic reflector, tracking system, receiver, thermoelectric converter, and power converter, as shown in Fig. 27.1. The parabolic reflector is applied to gather sunlight on the receiver. Then,

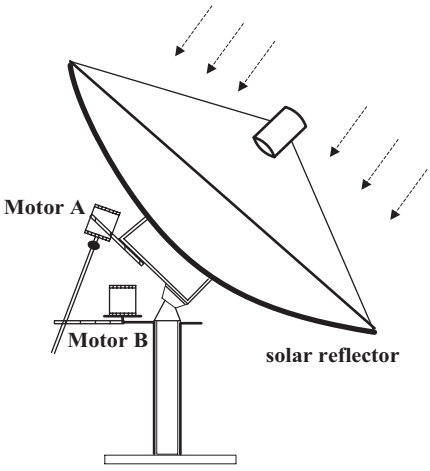


Fig. 27.1 Components of solar thermal power generation system

the heat is absorbed into the thermoelectric conversion system and converted into electrical energy [6].

27.2.1 Operation Principle and Composition of the System

As a core component, the tracking system usually uses the axis of condenser to parallel to the sun's rays. According to the location information of the sun, the biaxial mixed mode calculates the deflection angle. The servo motor drives the condenser to track, and sensors are used to correct the position deviation. The mixed mode reduces the tracking cumulative error. At the same time, the weather conditions can be avoided as much as possible.

However, the component aging and random disturbances such as wind loads may reduce the system's tracking accuracy and even make the system deviate from the steady state; therefore, the controllers which achieve rapid tracking and guarantee the tracking accuracy are not satisfied in practical application. On this basis, the external disturbance and the system uncertainty should be considered appropriately when designing the controller.

27.2.2 Model of Tracking Control System with Uncertainties

Figure 27.2 shows the subsystem block diagram of the tracking servo system. In Fig. 27.2, where f is the external random disturbance including the wind loads, θ_r is the input angle value, K is the amplification coefficient of magnetic amplifier, u is the input voltage of the servo motor, R and L are the resistance and inductance of the motor armature circuit, respectively, K_t is the torque coefficient of motor, T_e is the electromagnetic torque of motor, J_a is the inertia of the rotor and the motor load, B_a is the damping coefficient of the mechanical rotating system, K_v is the back electromotive force coefficient, ω is the motor angular velocity, θ_c is the positional signal of angle, and s is Laplace operator.

From the structure diagram in Fig. 27.2, θ_r denotes the system input, and θ_c is the system output. Three state variables $x1=\theta_c$, $x2=\omega$, *and* $x3=T_e$ are defined in the system. The wind loads can interfere with the servo motor load torque; therefore,

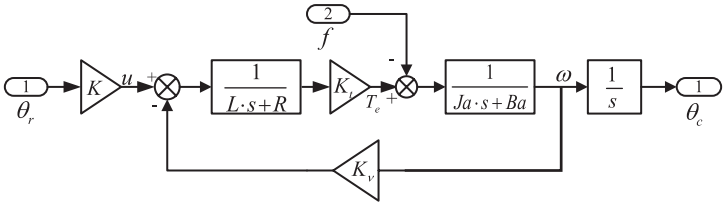


Fig. 27.2 Structure diagram of the tracking servo system

the wind loads disturbance f with variance $\sigma f2$ and zero mean white noise w is defined. Meanwhile, the nonlinearity and time-varying uncertainty of the system affect the tracking accuracy. Motor armature reaction may cause the nonlinearity of the torque coefficient K_t. Damping coefficient B_a varies within 20% range of the nominal value [7]. The matrix ΔA and ΔB containing system parameter uncertainties are added when the state space model of the system is established. Upon analysis, the state space model with parameter uncertainties and disturbance is obtained as follows:

$$\dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)u(t) + Gw(t), \tag{27.1}$$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -B_a / J_a & 1/J_a \\ 0 & -K_v K_t / L & -R/L \end{bmatrix}, \ \Delta A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\Delta B_a / J_a & 0 \\ 0 & -\Delta K_t K_v / L & 0 \end{bmatrix},$$

where $B = \begin{bmatrix} 0 & 0 & K_t \cdot K / L \end{bmatrix}^T$, $\Delta B = \begin{bmatrix} 0 & 0 & \Delta K_t \cdot K / L \end{bmatrix}^T$,

$G = \begin{bmatrix} 0 & -\sigma_f / J_a & 0 \end{bmatrix}^T$.

27.3 Tracking System Based on Robust Variance Control

Considering the system described by the state space model of Eq. (27.1), the RVC is a useful control algorithm to attain good control effect. Firstly, the matrixes ΔA and ΔB are transformed to the form as shown below [8]:

$$[\Delta A, \Delta B] = MF[H_1, H_2], \tag{27.2}$$

where $F \in R^{i \times j}$ satisfies the uncertainty matrix $FF^{\mathrm{T}} \leq I$, M, H_1, and H_2 are constant matrixes. In robust covariance control, the state feedback controller $u = K_f x$ is applied to the tracking system. Supposing $\tilde{A} = A + BK_f + MFH_1 + MFH_2K_f$, Eq. (27.1) is rewritten as:

$$\dot{x}(t) = \tilde{A}x(t) + Gw(t). \tag{27.3}$$

In order to ensure good transient performance of the system, all the poles of the closed loop system are located in the circular area with the center $(-q, 0)$ and the radius r on the complex plane. According to the principle of regional pole placement, the system of Eq. (27.3) meets the matrix inequality:

$$(A + qI)P(A + qI)^T - r^2 P + qGG^T < 0. \tag{27.4}$$

In order to guarantee that the system is asymptotically stable, a steady state variance matrix X exists and satisfies the following Lyapunov equation:

$$AX + XA^T + GG^T = 0. \tag{27.5}$$

Combining Eq. (27.4) with Eq. (27.5), the variance constraints of state variables are given:

$$[P]_{ii} < \sigma_i^2, (i = 1, 2, \cdots, n) \tag{27.6}$$

By configuring the system state variance less than a certain upper bound, the RVC makes the closed-loop system have good steady performance. Let $Y = K_f P$. With the Schur complement lemma, the solution to the feedback matrix K_f of RVC is converted to the solution to the linear matrix inequality.

For the given circular area *D(q, r)* and a set of constants σ_i^2 ($i = 1, 2, 3, \ldots, n$), there is a constant value ε ($\varepsilon > 0$), the matrix *Y,* and symmetric matrix *P* to make the following linear matrix inequality true [8]:

$$\begin{cases} \begin{bmatrix} -rP & PA^T + Y^T B^T + qP & PH_1^T + Y^T H_2^T \\ AP + BY + qP & qr^{-1}GG^T - rP + \varepsilon MM^T & 0 \\ H_1 P + H_2 Y & 0 & -\varepsilon I \end{bmatrix} < 0 \\ [P]_{ii} < \sigma_i^2, (i = 1, 2, \cdots, n) \end{cases}. \tag{27.7}$$

Based on the above analysis, the algorithm steps of the RVC are summarized as below:

1. Conduct the system model, including the disturbances matrix G, the parameters uncertainty matrixes ΔA and ΔB.
2. According to formula (2), the appropriate matrixes *M* and *F* are selected to get matrixes H_1 and H_2.
3. Considering the control requirement, σ_i^2 ($i = 1, 2, 3$) now are determined. Then, the linear matrix inequality (LMI) (3) is built up.
4. Applying the LMI toolbox to solve the LMI. Firstly, *Y, P,* and ε are defined using the "lmivar" function. Secondly, Eq. (27.7) is constructed by using the "lmiterm" function. In the end, the "feasp" function is used to solve *Y* and *P*.
5. The state feedback matrix is obtained as $K_f = YP^{-1} = [K_1, K_2, K_3]$.

27.4 Control Performance Analysis

Table 27.1 shows the basic parameters of the servo system [9].

ΔB_a and ΔK_t are set to ± 2e-5 and ± 0.009, respectively. The basic parameters from Table 27.1 are taken into Eq. (27.1):

Table 27.1 Basic parameters of the servo system

Parameter (unit)	Value	Parameter	Value
K	1	B_a	5e-3
L(H)	2.83e-3	K_t	0.0924
R(Ω)	1.75	K_v	0.093
$J_a(Kgm^2)$	3e-5		

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}0 & 1 & 0\\ 0 & -(166.7\pm 0.67) & 3.33e+4\\ 0 & -(3.04\pm 0.3) & -618.37\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$
$$+\begin{bmatrix}0\\ 0\\ 32.65\pm 3.18\end{bmatrix}\theta_r+\begin{bmatrix}0\\ (-3.33e+4)\sigma_f\\ 0\end{bmatrix}w. \tag{27.8}$$

In the uncertainty matrix of the Eq. (27.2), M is set to [0 1 0; 0 0 1]$^{\mathrm{T}}$ and F is set to the random number between −1 and 1. M, F, ΔA, and ΔB are put into the Eq. (27.2):

$$[H_1, H_2]=\begin{bmatrix}0 & -0.67 & 0 & 0\\ 0 & -0.3 & 0 & 3.18\end{bmatrix}. \tag{27.9}$$

After the model of the tracking servo system is established, the corresponding feedback controller should be analyzed. As mentioned above, all closed-loop poles are in the circular area with the center $(-q, 0)$ and the radius r on the complex plane. The number of the poles which meets the performance requirements would be reduced if q does not change and r gets smaller. Then, the solvability of the linear matrix inequality turns poor.

Assuming r is equal to q, the relationship between q and θ_c is analyzed. The randomness of θ_c should be constrained to be less than 0.1; therefore, $\sigma 1$=1e-4, $\sigma 2$=0.032, and $\sigma 3$=0.1. Because the coefficients of the noise input matrix G is too large, the wind loads with small variance constraints value would interfere with the tracking system to result in bad solvability. Thus, the standard deviation of the noise σf is set to 1e-5.

Figure 27.3 shows that the standard deviation value of the positional angle θ_c gets smaller as q gets larger. In Fig. 27.3, when $q<270$, the standard deviation value of θ_c rapid decreases. But it is still greater than the standard deviation value of the noise (σf is 1e-5). In other words, the designed controller has bad effects on noise reduction. When $q>270$, the standard deviation value of θ_c starts to be less than σf (σf is 1e-5). After q is greater than 2000, the standard deviation of θ_c tends to be stable. At this time, the standard deviation is near 2.35e-7. Thus, good control effect can be achieved when q is chosen to be greater than 270.

In Fig. 27.4, the variance value of the noise is greater as the variance value of output angle is greater. When q=2000, the variance value of the output angle changes slowly, and the noise is reduced well. When q=500 and the variance of the noise exceeds 5e-11, the variance value of the output angle starts to grow fast and the noise cannot be ignored.

Let q=1000, r=990, and the linear matrix inequalities (7) are solved. Then K_f=[−240.88, −0.88, −43.75]. The contrast between output angle and the noise is shown in Fig. 27.5. And the standard deviation value of the output angle is 1.2e-6, and the noise (σf is 1e-5) is suppressed well.

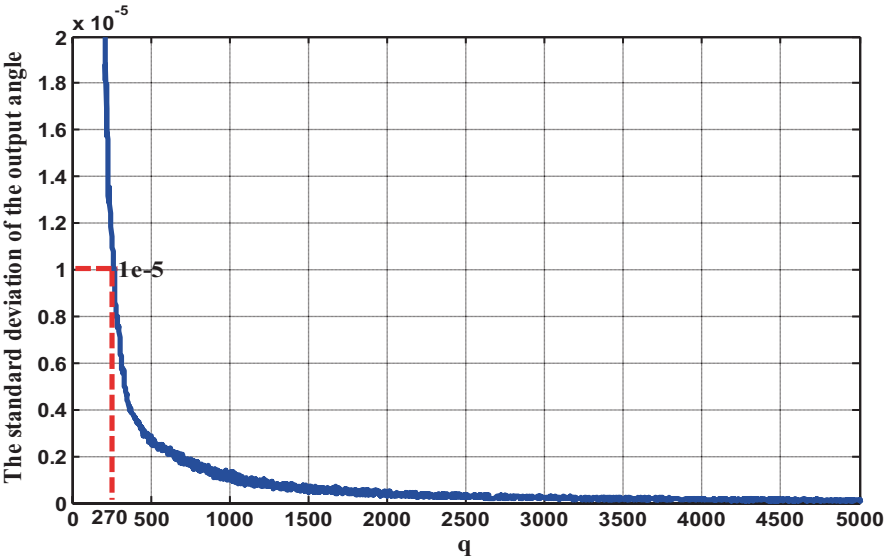


Fig. 27.3 Relationship between q and standard deviation of output angle

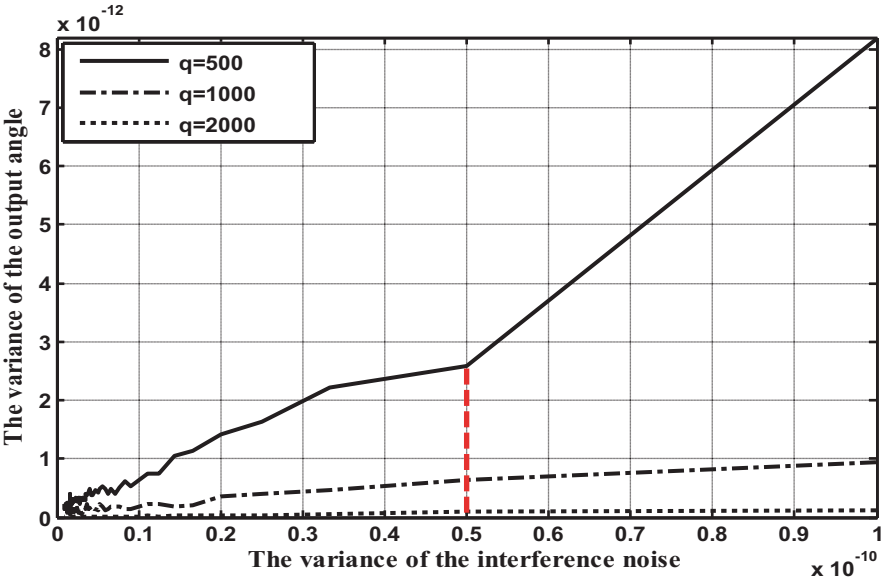


Fig. 27.4 Relationship between the noise variance and the output angle variance

The contrast of the unit step response between proportional integral derivative (PID) control (K_p=88, K_i=2, K_d=0.14) [3] based on the stability boundary method and RVC is shown in Table 27.2. From Table 27.2, the results show that the RVC has a better transient performance than a traditional controller.

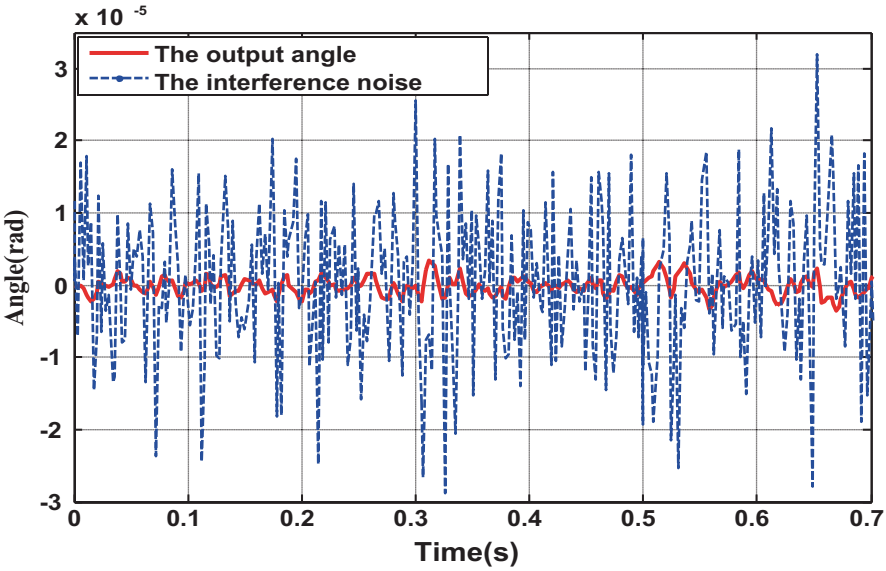


Fig. 27.5 Contrast between the output angle and the noise

Table 27.2 Comparison of transient performance of control methods

Control methods	Overshoot (%)	Rise time(s)	Regulation time(s)
PID	17.6	0.007	0.0244
RVC (q=1000)	0	0.006	0.0105
RVC (q=5000)	0	0.0048	0.0088

PID proportional integral derivative, *RVC* robust variance control

27.5 Conclusion

In this chapter, the state space model of dish solar generation tracking servo system with the random disturbances of wind load and system parameter uncertainties is built. Then, the robust variance constraints algorithm for tracking system is given to overcome the disturbances and uncertainties; furthermore, the influence of circle centre location on control performance is analyzed. The simulation results show the RVC is an effective control strategy to be applied in the solar tracking systems, thus offering the possibility of an advanced control application in the renewable power generation process.

References

1. Bhamidipati UK. Review of renewable solar energy. Sacramento: California State University; 2012.
2. Li YQ, Liao SM, Rao ZH, Liu G. A dynamic assessment based feasibility study of concentrating solar power in China. Renew Energy. 2014;69(2014):34–42.
3. Peng CQ, Peng YD, Hu YP, Yin X, Hu L. Research and implementation of tracking control system for a high-power single-disc solar concentrator. J Hunan Univ of Sci Technol. 2012;27(3):28–34. (In Chinese)
4. Alata M, AI-Nimr MA, Qaroush Y. Developing a multipurpose sun tracking system using fuzzy control. Energy Convers Manage. 2005;46(7):1229–45.
5. ALQahtani AH, Abuhamdeh MS, Alsmadi YM, Utkin VI. Photovoltaic power optimization using sliding mode control with a two-axis tracking system. Energy tech, 2013 IEEE; Cleveland, OH: IEEE; 2013. p. 1–6.
6. Abbas M, Boumeddane B, Said N, Chikouche A. Dish Stirling technology: A 100 MW solar power plant using hydrogen for Algeria. Int J Hydrog Energy. 2011;36(7):4305–14.
7. Fang YM, Wang L, Wang YQ, Jiao XH. The research of speed robust tracking control on DC motor considering the nonlinear of armature reaction. Proc CSEE. 2003;23(5):136–9. (In Chinese)
8. Yu L. Robust control of uncertain systems with regional pole and variance constraints. Acta Automatica Sin. 2000;26(4):509–14. (In Chinese)
9. Kong XC, Xiao H, Chen LG, Sun FR. Simulation of fuzzy PID controls in light source tracking servo-system. Computer Simul. 2006;23(1):152–4, 164. (In Chinese)

Part III
Power Automation

Chapter 28
Internal AC Fault of a Converter Station Based on Modular Multilevel Converter High Voltage Direct Current

Tianfu Shen and Wenbo Shi

Abstract The chapter introduces the structure and operational mechanism of a modular multilevel converter based on high voltage direct current (MMC-HVDC), analyzes the internal alternating current (AC) system fault with emphasis, and puts forward the corresponding protection and control strategy. The MMC-HVDC connected to a wind farm simulation model is set up in power system computer aided design/electromagnetic transient including direct current (PSCAD/EMTDC). The simulation results show that during the interphase, the short circuit fault in the AC bus could cause serious accidents and the fault could be passed through the DC line to the nonfault converter. So it often requires to shut down the converter station—even both. However, the AC voltage on the nonfault converter would not be impacted. When troubleshooting, the system could adjust by self-tuning the control system into a stable state again.

Keywords Modular multilevel converter · HVDC · Internal AC system fault

28.1 Introduction

The high voltage direct current (HVDC) Flexible is a new type of high voltage direct current transmission technology, which is based on voltage source converter (VSC), a full-controlled device insulated gate bipolar transistor (IGBT) and pulse width modulation (PWM) technology [1]. The modular multilevel converter (MMC) with scalable and modular design features, can effectively achieve the power conversion on a high voltage level, which is the most popular topology applied to VSC-HVDC converter currently. At present, though there are some HVDC Flexible projects that have already being put into operation, the HVDC Flexible technology research mainly focused on the study of topology and control strategy. However, for the fault analysis and corresponding overvoltage protection studies, few articles are

T. Shen (✉) · W. Shi
Electric Power of South China University of Technology,
Tianhe, Guangzhou, 510640 Guangdong, China
e-mail: 78474665@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_28

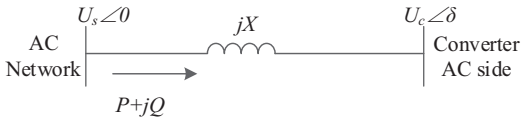


Fig. 28.1 Power flow between the converter and alternating current (AC) network diagram

introduced currently. Based on the above situation, the chapter describes the basic structure and operating principle of MMC, analyzed internal AC system fault simulation in MMC, and discussed corresponding protection strategies.

28.2 MMC Basic Structure and Operating Principle

28.2.1 Power Transmission Principle

The VSC-HVDC power transmission is shown in Fig. 28.1. When the commutation reactor loss and harmonic components are ignored, the active power P and reactive power Q transfer between the VSC and AC network are as shown below [2]:

$$P = -\frac{U_s \cdot U_c}{X} \sin\delta \tag{28.1}$$

$$Q = \frac{U_s \cdot (U_s - U_c \cos\delta)}{X} \tag{28.2}$$

U_s represents the fundamental voltage component in the common connection point of AC bus, U_c represents the fundamental voltage component of VSC output, δ is the phase difference between U_s and U_c, and X is the equivalent reactance. By the formulas (1) and (2), the active power transmission depends on phase shift angle δ, and the reactive power mainly depends on the fundamental voltage amplitude U_c in the AC output of VSC.

28.2.2 MMC Working Mechanism

Nowadays, the HVDC Flexible projects mostly adopt the MMC topology, which uses half-bridge series technology. Its outer characteristic has the VSC feature and can do the four-quadrant operation. As shown in Fig. 28.2a, a three-phase MMC main circuit topology structure contains six bridge arms; each arm consists of n number of submodule (SM) units and one reactor in series [3]. The SM unit structure as shown in Fig. 28.2b consists of an insulated gate bipolar transistor (IGBT) as the switch element and a half-bridge DC energy storage capacitor C in parallel.

The IGBTs of up-bridge and down-bridge arm cannot conduct at the same time in SM unit, thus there are three kinds of working state as shown in Fig. 28.3.

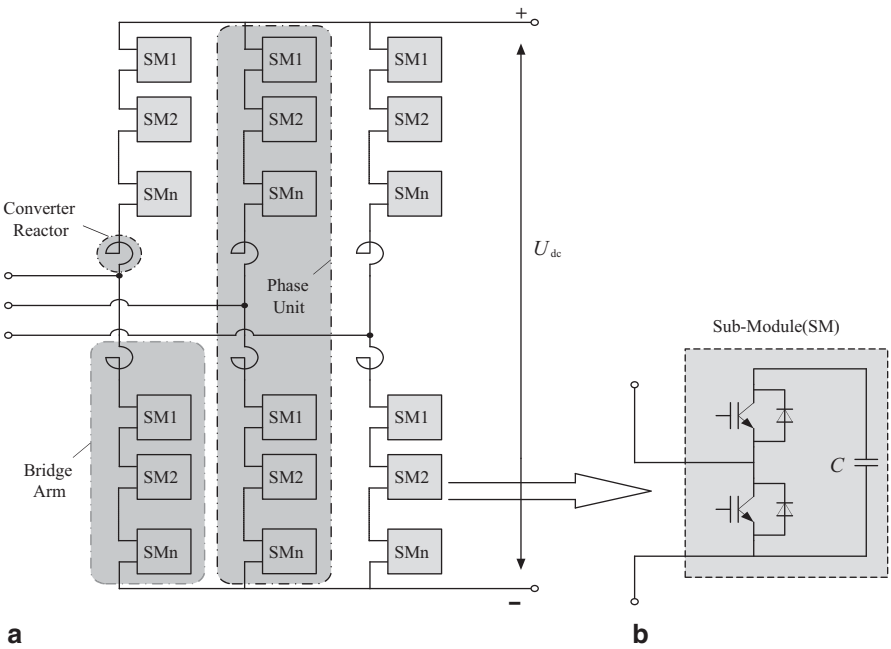


Fig. 28.2 Three-phase modular multilevel converter (*MMC*) main circuit topology diagram

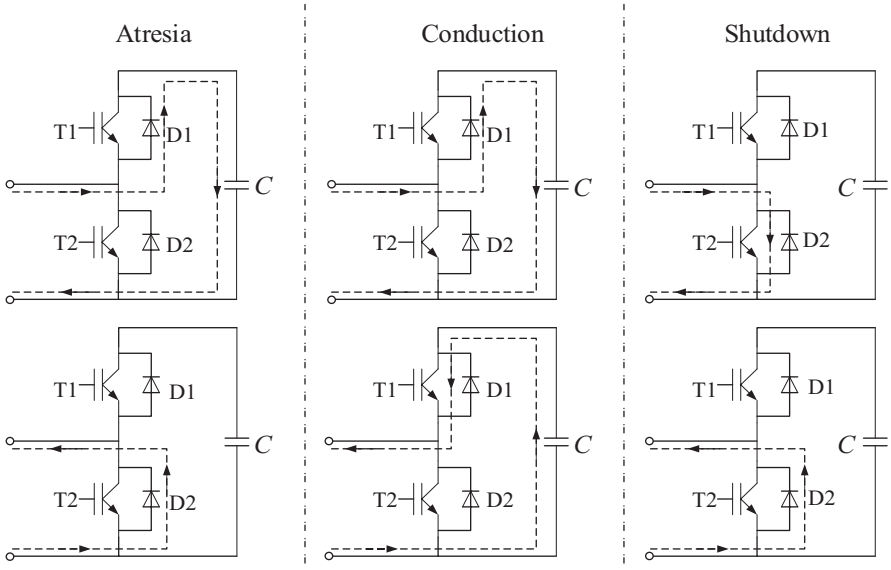


Fig. 28.3 The mechanism of the modular multilevel converter (*MMC*) submodule (*SM*) unit

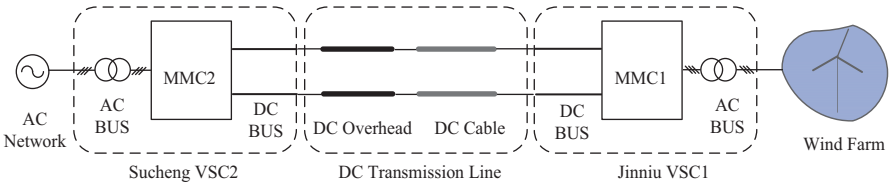


Fig. 28.4 Double-end modular multilevel converter based on high voltage direct current (*MMC-HVDC*) system topology of Nanao wind farm. *DC* direct current, *AC* alternating current

(1) Atresia: upper and lower IGBT (T1, T2) are in the off state, the forward continuity of the anti-parallel diode (D1, D2) decides the SM state. When the current flows through diode D1, the bridge arm is in charged condition, and the capacitor C connected in series; when the current flows through diode D2, the capacitor C is bypassed. (2) Conduction: T1 is opened, T2 turned off. Regardless of the direction of the current, the SM output voltage is the capacitor voltage, and the direction of the current determines whether the capacitor is charging or discharging. (3) Shutdown: T1 is off, T2 opened; the current follows through T2 or D2; the SM capacitor is always bypassed, and therefore the output voltage of the module is 0.

28.3 Fault Simulation

The MMC-HVDC system mainly consists of converter station, DC transmission line, AC power grid, and wind farm (or passive load) [4]. The internal MMC mainly contains DC bus, AC bus, and the control and protection equipment. The DC transmission line is composed of DC overhead lines or DC cables, or both. In this chapter, we take Shantou Nanao's wind farm double-end HVDC Flexible project for an example, as shown in Fig. 28.4. And AC system fault in the internal converter shows in Fig. 28.5.

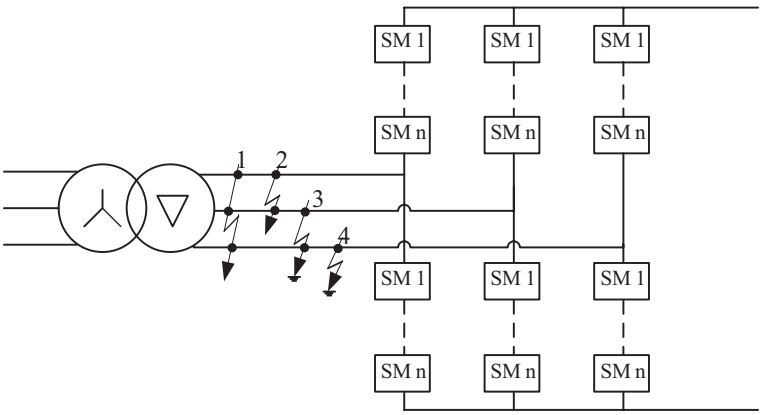


Fig. 28.5 Alternating current (*AC*) system fault in the internal converter. *SM* submodule

28.3.1 AC Bus Fault

AC bus fault refers to a short circuit fault that occurs between the AC transformer secondary side and the AC output of VSC side, as shown in Fig. 28.5 in which the fault point is between 1 and 4. According to the different fault set points, the fault conditions can be divided into interphase short circuit and single-phase ground short circuit (corresponding to the fault 4 in Fig. 28.5). The interphase short circuit contains a three-phase ground short circuit, a two-phase short circuit, and a two-phase ground short circuit fault [5], respectively corresponding to the fault conditions between 1 and 3 as in Fig. 28.5. Choose the most serious fault three-phase short circuit in the AC bus as an example to simulate and put forward the protection methods. The other fault characteristics are similar to the three-phase short circuit fault, only the fault level is a little weaker.

28.3.2 Fault Simulation Waveforms

The MMC-HVDC system capacity is 200 MW and the rated DC voltage is ± 160 kV. The MMC-rated AC output voltage is 110 kV, which is connected through a double-winding transformer with the 220 kV AC system. The VSC2 adopts the constant DC and AC voltage control strategy, and the VSC1 uses the constant active power and the AC voltage control strategy. Each bridge arm uses 10 SMs in the simulation model. Setting three-phase nonmetallic ground fault besides the AC side of the wind farm, the fault point seated in VSC1 side. When the system accessed the steady state, the input of the fault is at 0.5 s and the fault duration is 0.2 s. Fault features are shown in Figs. 28.6 and Fig. 28.7.

28.3.3 Failure Analysis

According to the above waveforms, after the three-phase short circuit occurred, the system voltage instantaneously dropped, but the AC voltage of the VSC fault side had a certain inertia in the adjustment process that could not follow the system voltage drop quickly; so the AC side overcurrent occurs. Meanwhile, the fault-side VSC simultaneously transmittes active power P and reactive power Q, and the current of DC side I_{dc} would decline. However, the nonfault-side converter transmitted power ability was not affected. The input and output energy imbalance would be reflected in the total DC voltage U_{dc} and the capacitor voltage of SMs. Once the DC voltage could not be maintained at the set value, it would affect the operating characteristics of nonfault-end, and the fault spreads from the converter station at one side to the other side.

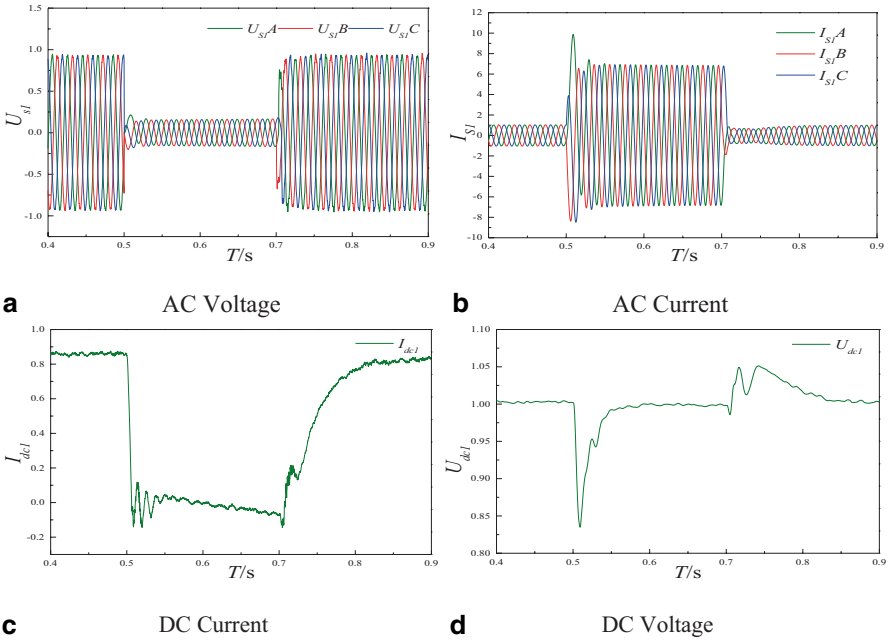


Fig. 28.6 The fault waveforms of voltage source converter 1 (*VSC1*) side. *AC* alternating current, *DC* direct current

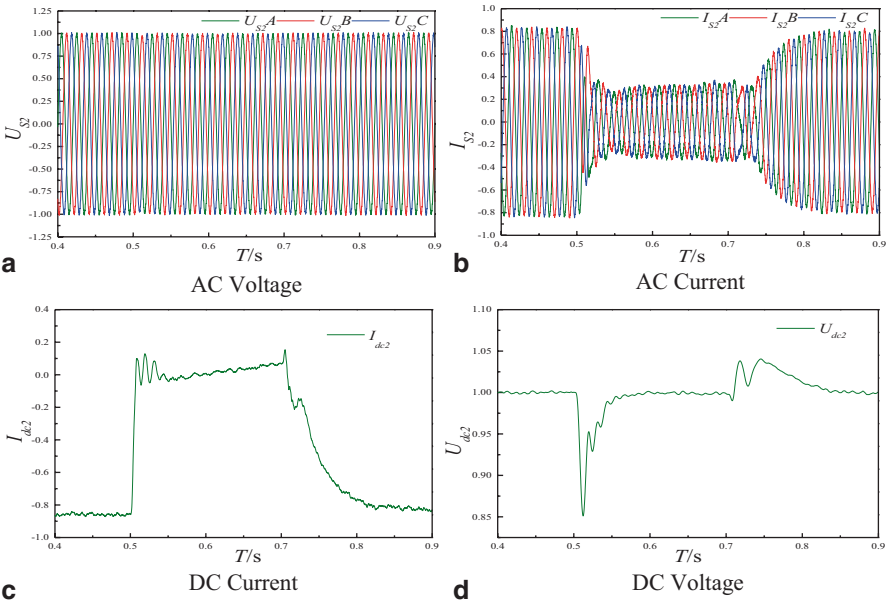


Fig. 28.7 The fault waveforms of voltage source converter 2 (*VSC2*) side. *AC* alternating current, *DC* direct current

28.4 Protection and Control

The AC bus interphase short circuit is mostly caused by defective insulation [6], and due to the fact that the electrical distance between the fault point and the converter is small, the fault is always serious. By the analysis of fault waveforms it was shown that when the system AC voltage sagged, the AC current would rise, and the more severely voltage sagged, the more obviously the AC current rose. The bridge arm current overflowed due to the increasing AC component in the converter; the power input was blocked; the DC current that followed decreased, and the fault passed through the fluctuation of DC line to the other side of the VSC, but the AC voltage on the other side would not be impacted. While troubleshooting, the system could adjust by a self-tuning control system into a stable state again.

Through the above analysis, on the one hand, ensuring the interior of the converter does not lead to the overvoltage or overcurrent, and the security and integrity of all kinds of devices are maintained. On the other hand, maintaining the total DC voltage stable prevents the spread of the fault. So often it requires shutting down the HVDC Flexible system at one side of the converter station, sometimes even on both. So the protection method as shown in Fig. 28.8 is as follows: (1) detect whether the arm current is overcurrent or not; (2) if so, the bridge arm overcurrent caused the overcurrent protection action, blocking the fault converter. (3) Detect whether the DC voltage is abnormal or not; (4) if so, the DC undervoltage or overvoltage occurs, coordinating the control and protection of the nonfault side.

If the fault VSC adopts constant DC voltage control, it will lose the ability to control DC voltage after atresia. This energy imbalance can lead to too high or too low DC voltage, causing the DC-side overvoltage/undervoltage; adopting the protection action and locking the nonfault converter to protect the converter capacitor element, the entire MMC-HVDC system stops running. If the wind farm is still connected to the MMC-HVDC system, it will conduct uncontrolled rectifier through the diode, continuing injecting energy into the MMC-HVDC system. Therefore, the wind farms must also be removed.

If the fault converter adopts fixed AC voltage control, the constant DC voltage control loop of the nonfault converter will gradually adjust the input/output of active power after locking, maintaining one constant DC voltage, and the fault will not spread. However, as the existence of the fault point and the voltage of the wind farm grid interface reduced, the fan needs the low voltage control.

28.5 Conclusion

1. The characteristics of interphase short circuit fault in the AC bus: (1) AC bus voltage drop; (2) AC side and bridge arm current surge; (3) DC side current fall; (4) transmission power down; (5) input and output energy imbalance can cause

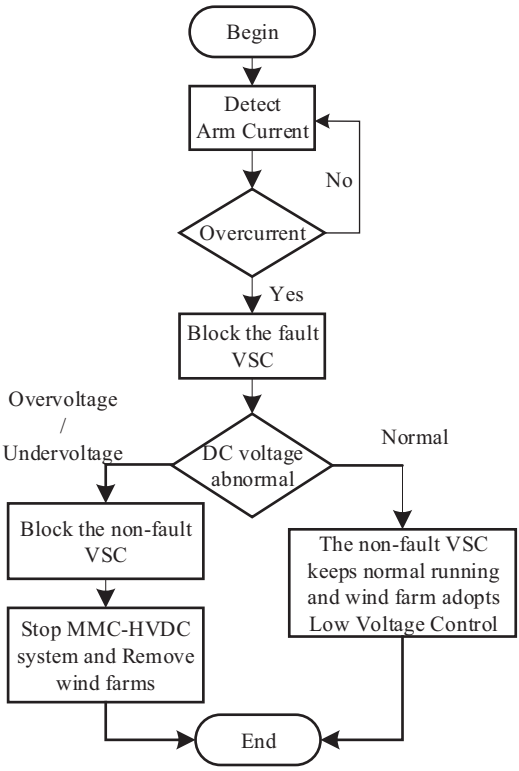


Fig. 28.8 Protection method for the interphase short circuit fault in alternating current (*AC*) bus. *DC* direct current, *VSC* voltage source converter, *MMC-HVDC* modular multilevel converter based on high voltage direct current

DC system voltage and SM capacitor voltage abnormal; (6) the abnormal voltage of the DC system leads to the spread of fault.

2. The interphase short circuit fault occurred in VSC usually belongs to permanent fault; this fault can be passed through the DC line to the other side of VSC, which makes the entire line of the MMC-HVDC power fluctuate and run instable. It is a serious fault that should lead to timely shutting down of the power and electronic device and disconnecting the grid circuit breaker.
3. Due to the electrical distance between the fault point and the converter is small, the fault is relatively serious, and in order to ensure the security and integrity of all kinds of devices, and to prevent the spread of fault, so often it is needed to shut down the MMC-HVDC system at one side of the converter station, even both.

References

1. Schettler F, Huang H, Christl N. HVDC transmission systems using voltage sourced converters design and applications. Proceedings of 2000 IEEE Power Engineering Society Summer Meeting, IEEE; 2000. p. 715–20.
2. Wang H, Tang G, He Z, Yang J. Efficient grounding for modular multilevel HVDC converters (MMC) on the AC side. IEEE Trans Power Deliv. 2014;29(3):1262–72.
3. Li X, Song Q, Liu W, Rao H, Xu S, Li L. Protection of nonpermanent faults on DC overhead lines in MMC-based HVDC systems. IEEE Trans Power Deliv. 2013;28(1):483–90.
4. Flourentzou N, Agelidis V, Demetriades G. VSC-based HVDC power transmission systems: an overview. IEEE Trans Power Electr. 2009;24(3):592–602.
5. Zhou H, Shen Y, Li M. Research on insulation coordination for converter stations of Zhoushan multi-terminal VSC-HVDC transmission project. Power Syst Technol. 2013;37(4):879–90.
6. Liu Z, Song Q, Liu W. VSC-HVDC system based on modular multilevel converters. Automation Electr Power Syst. 2010;34(2):53–8.

Chapter 29
Harmonic Analyzing Based on Cubic Spline Interpolated Arithmetic Fourier Transform

Jie Wu, Kaipei Liu, Jian Le, Li Wang and Yihuang Chen

Abstract The data sampling frequency of the signals in digital substation is generally fixed to 10 kHz. The fast Fourier transform may not be suitable for its harmonic analysis. The arithmetic Fourier transform (AFT) is more appropriate for analyzing discrete signals due to the advantages such as simpler computation, better parallelism, and no limitation on the number of sampling points. It requires non-uniform sampling points, so the uniform sampled signals should be interpolated when using AFT. The zero interpolation is the most widely used method of AFT. It produces a negative effect on the accuracy of the harmonic analysis, which cannot satisfy the requirements of the power system. This chapter proposes a new interpolation method of AFT after comparing the accuracy performance of four interpolation methods, i.e., the zero interpolation, the first-order linear interpolation, the piecewise cubic hermite interpolation, and the cubic spline interpolation. Finally, the cubic spline interpolation is selected to improve the accuracy due to its higher precision and better stability. The MATLAB simulation results show that the new interpolation can meet the requirements of power system harmonic analysis, make AFT computational more accurate, and provide new ways for harmonic analysis.

Keywords Arithmetic Fourier transform · Interpolation methods · Harmonic analysis · Precision

J. Wu (✉) · K. Liu · J. Le · L. Wang · Y. Chen
School of Electrical Engineering, Wuhan University, 430072 Wuhan, China
e-mail: wjwd2008@whu.edu.cn

K. Liu
e-mail: kpliu@whu.edu.cn

J. Le
e-mail: lej01@mails.tsinghua.edu.cn

L. Wang
e-mail: 417197078@qq.com

Y. Chen
e-mail: 657825784@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_29

29.1 Introduction

At present, harmonic analysis in power systems is mostly based on fast Fourier transform (FFT) and discrete wavelet transform [1–4], while the calculation process of FFT harmonic analysis is complicated and contains lots of child processes [5]. The filter used in wavelet transform has serious influence on the analysis results [6]. The arithmetic Fourier transform (AFT) has simpler construction and less multiplication, which is more suitable for harmonic analysis. Nonuniform sampling is required in AFT, while the power system only has uniform sampling; so the discrete signals should be continued by interpolation methods before nonuniform sampling is done. The zero interpolation method AFT used widely [7] in processes will bring huge errors to analysis results and influence the precision of harmonic analysis badly, which cannot meet the requirements of the power system.

In this chapter, the cubic spline interpolation as a new interpolation method for AFT is found to solve the problems above. To prove its superiority, the principle of the zero interpolation, the first-order linear interpolation, the piecewise cubic hermite interpolation, and the cubic spline interpolation are introduced to make comparison. Then, MATLAB is used to achieve the simulations of the signal, which contains harmonics with the order up to 20. Finally, the cubic spline interpolation is chosen as the new method according to the results.

29.2 Arithmetic Fourier Transform

Arithmetic Fourier transform (AFT) is a method to calculate the Fourier coefficients of periodic signals by introducing the Mobius inversion formula [8]. It was named by Tufts and Sadasiv in 1988.

Comparing AFT with FFT, the advantage of AFT is that the numbers of its multiply operations can be reduced markedly by replacing multiplication with moderate addition operation. For example, in N-point FFT, due to its symmetry, the number of plural multiply operation and plural addition operation is $N/2 \times \log_2 N$ and $N \times \log_2 N$, respectively. While for AFT, these two numbers are reduced to N/2 and N^2. Because the efficiency performed by the computer of addition operation is faster than multiply, the computation speed of AFT is faster than FFT. AFT also has the better parallelism, and it can be used to calculate Fourier coefficients of complex continuous functions and is especially suitable for VLSI design.

The samples of a signal, for which the fundamental frequency is $f = 1/T(\text{Hz})$, can be expressed as $A(t)$.

$$A(t) = a_0 + \sum_{n=1}^{N} a_n \cos 2\pi nft + \sum_{n=1}^{N} b_n \sin 2\pi nft \tag{29.1}$$

Use AFT to analyze the harmonics of this signal, where a_0, a_n, and b_n can be calculated as follows:

$$a_0 = \frac{1}{T}\int_0^T A(t)dt, \tag{29.2}$$

$$a_n = \sum_{l=1,3,5,\ldots}^{[N/n]} u(l)B(2nl,0), \tag{29.3}$$

$$b_n = \sum_{l=1,3,5,\ldots}^{[N/n]} u(l)(-1)^{(l-1)/2} B(2nl,1/4nl), \tag{29.4}$$

where

$$B(2n,\alpha) = \frac{1}{2n}\sum_{m=0}^{2n-1}(-1)^m A(\frac{m}{2n}T + \alpha T), \tag{29.5}$$

where $m = 0,1,2,\ldots,2n-1$, and $-1 < \alpha < 1$;

The Mobius function is defined as follows:

$$u(l) = \begin{cases} 1, l = 1 \\ (-1)^r, l = p_1 p_2 \ldots . p_r, \\ 0, \exists p, makes\ p^2 \mid l \end{cases} \tag{29.6}$$

where p_i represents different prime numbers. When l is the prime number or the product of odd prime numbers, $u(l) = -1$. When l is the product of even prime numbers, $u(l) = 1$. When l has square factors, $u(l) = 0$.

The primary method used in AFT is zero interpolation. While its accuracy is not high enough to meet the requirements of the power system, other interpolation ways are found to improve the accuracy.

29.3 Interpolation Methods

Assume that the interpolation is in the interval [a, b], where $a = x_0 < x_1 < \ldots < x_n = b$, the corresponding function values are y_0, $y_1 \ldots$ y_n, and $A(t)$ is the interpolation function.

29.3.1 Zero Interpolation and the First-Order Linear Interpolation

The principle of the zero interpolation is shown in Fig. 29.1a. It means that if the points needed by AFT have not been sampled, then they will use the value of the nearest point to replace.

The first-order linear interpolation connects the adjacent sampled points with straight lines. If the points needed by AFT have not been sampled, their values can be obtained by the interpolation function curve, as it is shown in Fig. 29.1b.

In Fig. 29.1, points marked with "*" represent the original sample points, horizontal coordinate x means the points that have not been sampled, and the ordinate axis y represents the corresponding values of the points.

29.3.2 The Piecewise Cubic Hermite Interpolation

The hermite interpolation is the basis for spline interpolation and has strong adaptation and controllability [9–10], as shown in Fig. 29.2.

It requires that the interpolation function $A(t)$ is the three-order polynomial in the child interpolation interval $[x_{i-1}, x_i (i=1,\ldots,n)]$ and at the same time meets the conditions as follows: $A(x_i)=y_i$, $A'(x_i)=y_i'$. The way to calculate $A(x_i)$ can be described as follows (29.7):

$$\begin{cases} A(x_i) = y_{i-1} + a_{i1}(x - x_{i-1}) + a_{i2}(x - x_{i-1})^2 + a_{i3}(x - x_{i-1})^3 \\ a_{i1} = A'(x_{i-1}) \\ a_{i2} = \dfrac{3(y_i - y_{i-1})}{m_{i-1}^2} - \dfrac{2A'(x_{i-1}) + A'(x_i)}{m_{i-1}} \\ a_{i3} = -\dfrac{2(y_i - y_{i-1})}{m_{i-1}^3} + \dfrac{A'(x_{i-1}) + A'(x_i)}{m_{i-1}^2} \end{cases}, \tag{29.7}$$

where $m_{i-1}=x_i - x_{i-1}$.

For the two end nodes, the first derivative $A'(x_i)$ can be calculated as follows:

$$\begin{cases} A'(x_0) = \dfrac{x_1 - x_0}{y_1 - y_0} \\ A'(x_n) = \dfrac{x_n - x_{n-1}}{y_n - y_{n-1}} \end{cases}, \tag{29.8}$$

while for other points, the first derivative $A'(x_i)$ can be calculated as follows:

$$A'(x_i) = \frac{3(y_{i+1} - y_i)(y_i - y_{i-1})(x_{i+1} - x_{i-1})}{(y_{i+1} - y_i)(-x_i^2 + x_{i-1}^2 + 2x_{i+1}x_i) + (y_i - y_{i-1})(-x_{i+1}^2 - x_i^2 + 2x_i x_{i-1})} \tag{29.9}$$

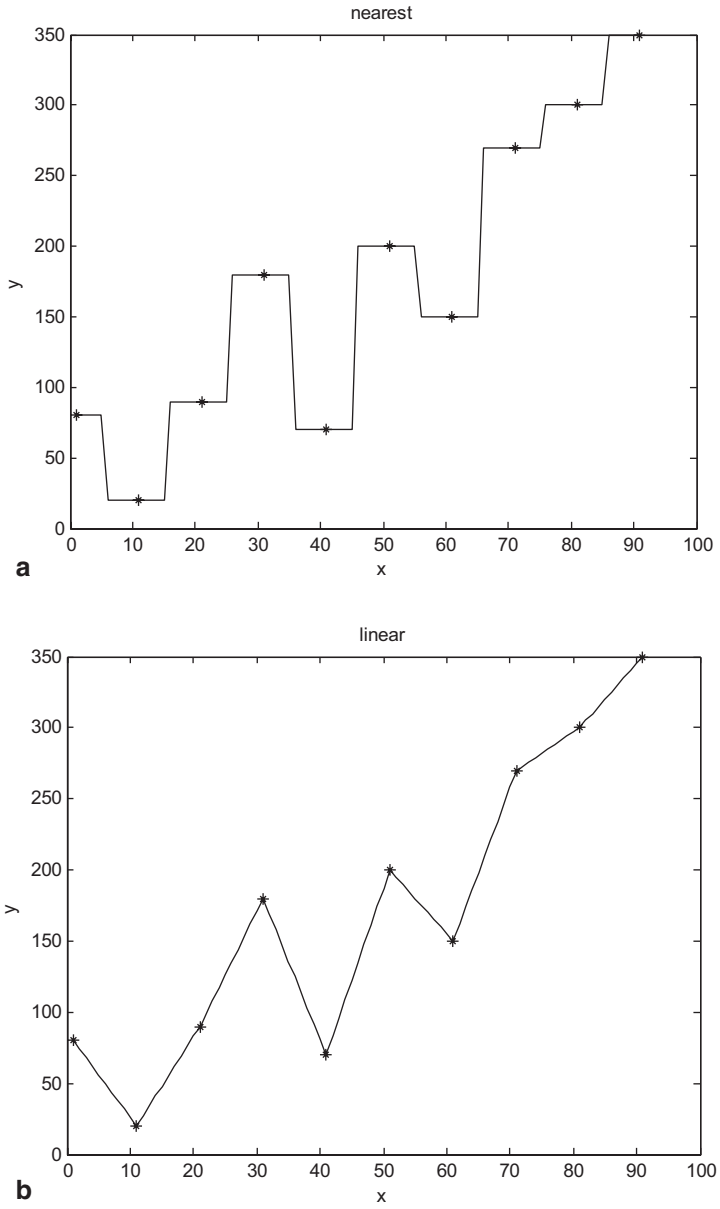


Fig. 29.1 **a** Zero interpolation. **b** First-order linear interpolation

The interpolation function $A(t)$ can then be calculated by substituting $A'(x_i)$ in (29.7).

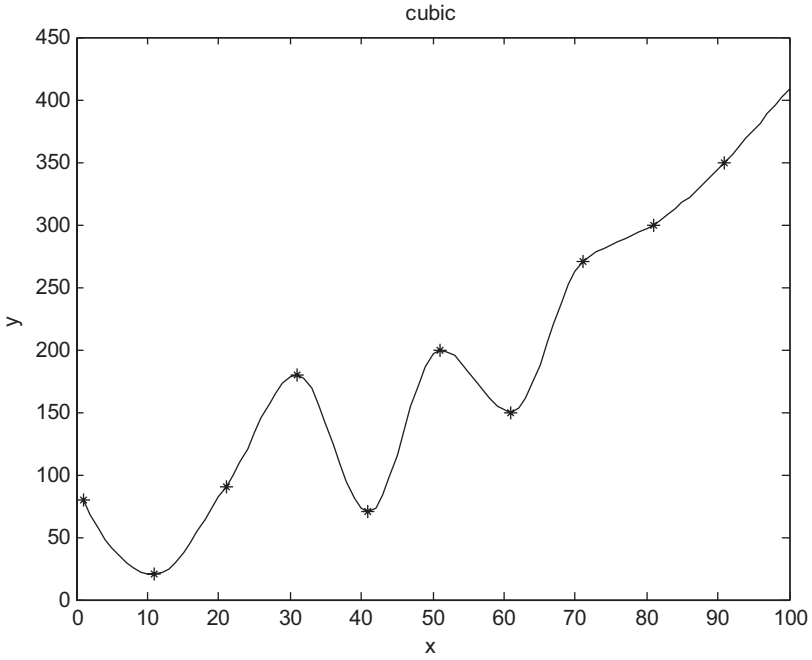


Fig. 29.2 Piecewise cubic hermite interpolation

29.3.3 The Cubic Spline Interpolation

The cubic spline interpolation connects the adjacent sampled points with no more than three orders of polynomial, as shown in Fig. 29.3. It requires $A(x_i)=y_i$, and the interpolation function $A(t)$, the first derivative $A'(x)$, and the second derivative $A''(x)$ are continuous functions in the child interval [11], and $A(x)$ can be calculated as follows:

$$A(x) = A'(x_{i-1})\frac{(x_i - x)^3}{6m_{i-1}} + A'(x_i)\frac{(x - x_{i-1})^3}{6m_{i-1}} + (y_{i-1} - \frac{A'(x_{i-1})m_{i-1}^{\ 2}}{6})\frac{x_i - x}{m_{i-1}} + (y_i - \frac{A'(x_i)m_{i-1}^{\ 2}}{6})\frac{x - x_{i-1}}{m_{i-1}} \tag{29.10}$$

where $m_{i-1}=x_i-x_{i-1}$. For $A(x)$ which has continuous second derivative in point x_i, add free boundary conditions as follows:

$$A''(x_0) = y_0'' = 0$$
$$A''(x_n) = y_n'' = 0 \tag{29.11}$$

The equation of A'(x) can be obtained as follows:

$$\mu_i A'(x_{i-1}) + 2A'(x_i) + \lambda_i A'(x_{i+1}) = \delta_i \tag{29.12}$$

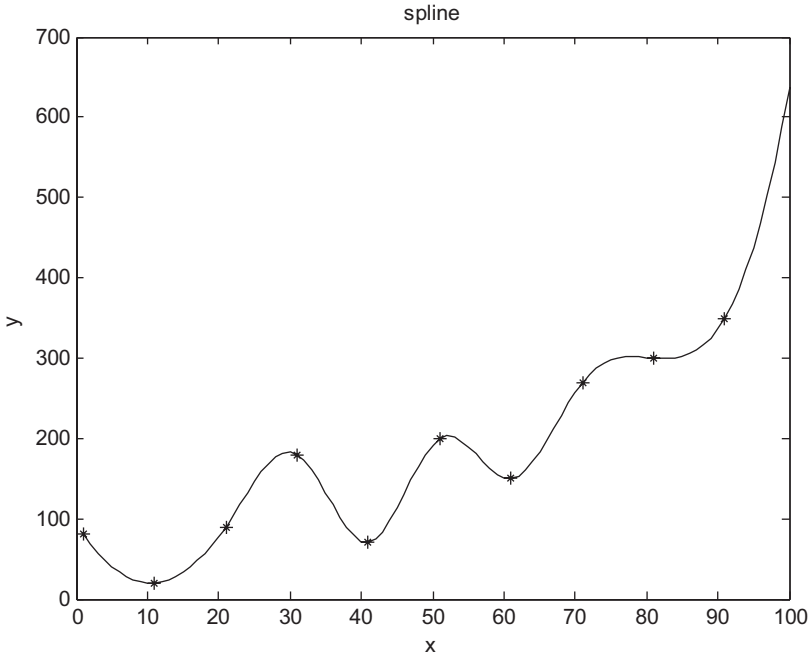


Fig. 29.3 Cubic spline interpolation

where

$$\begin{cases} \mu_i = \dfrac{m_{i-1}}{m_{i-1}+m_i} \\ \lambda_i = 1-\mu_i \\ \delta_i = 6\left(\dfrac{y_{i+1}-y_i}{m_i} - \dfrac{y_i - y_{i-1}}{m_{i-1}}\right)(m_i + m_{i-1})^{-1} \end{cases} \tag{29.13}$$

where $\mu_0 = 0, \delta_0 = \delta_n = 0, \lambda_n = 0.$

To solve the Eq. (29.11)–(29.13), bring the results in (29.10), the cubic spline interpolation in every child interpolation interval can be obtained.

29.4 Simulation Analysis

The simulation software is MATLAB. Write the computer programs in its function interface according to the principles above and the results can be obtained.

The procedure of the simulation includes signal sampling, four methods interpolation, sampling of the interpolated functions, and using AFT to analyze signal and error comparison, all of which are shown in Fig. 29.4.

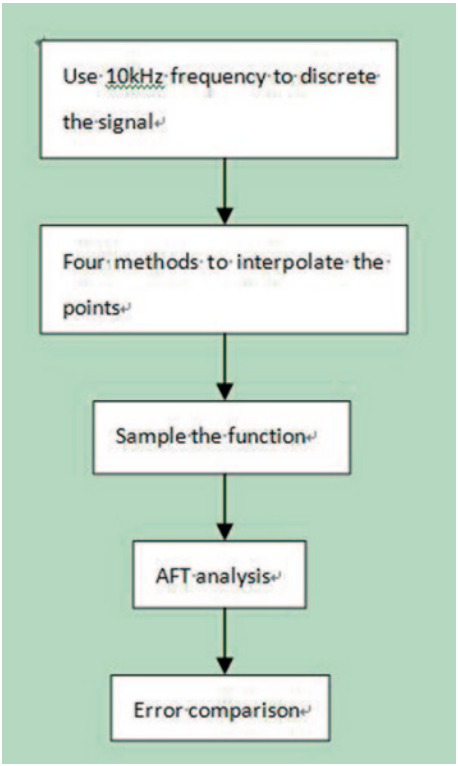


Fig. 29.4 The procedure of the simulation

A distortion signal $f(x)$ contains harmonics with the order up to 20 as shown in (29.14):

$$f(x) = 10\sin(2\pi f_0 t) + \sum_{k=2}^{20} \sin(2\pi k f_0 t) \tag{29.14}$$

where k is the harmonic order, and the fundamental frequency is $f_0 = 50$ Hz. Sample the signal shown by (29.14) with the rate of 10 kHz, and then interpolate the discrete sampled points into continuous function $A(t)$ by the four methods introduced above. Analyze the frequency spectrum characteristics of $A(t)$ by AFT, and then the amplitudes of each harmonic component can be obtained. Compare the result with the amplitudes of function $f(x)$, work out the errors.

After using the zero interpolation, the first-order linear interpolation, the piecewise cubic hermite interpolation, and the cubic spline interpolation, the frequency spectrum characteristics analyzed by AFT are shown in Table 29.1. The relative error comparison of these four interpolation methods is shown in Table 29.2. Fig. 29.5 shows the calculation of the relative error curves. It can be seen that the calcula-

Table 29.1 Results of AFT analysis after interpolation

Harmonic times		1	2	3	4	5	6	7	8	9	10
Amplitude (A)	Zero	10.208	0.910	0.795	0.795	0.906	1.011	1.084	1.070	0.951	1.059
	First order	9.988	0.987	0.991	1.018	0.985	1.004	0.990	1.001	0.993	0.997
	Hermite	10.031	0.988	1.007	1.002	0.984	1.006	1.008	1.000	1.000	1.001
	Spline	10.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Harmonic times		11	12	13	14	15	16	17	18	19	20
Amplitude (A)	Zero	1.009	0.931	0.981	0.922	0.944	0.960	1.039	0.944	1.025	1.145
	First order	0.983	0.986	0.986	0.975	0.982	0.980	0.981	0.973	0.969	0.961
	Hermite	1.000	0.999	0.999	0.993	0.990	0.992	1.004	1.005	1.002	1.004
	Spline	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 29.2 Comparison of the relative error

Harmonic times		1	2	3	4	5	6	7	8	9	10
Relative error (%)	Zero	2.08	9.0	20.5	15.3	9.4	1.1	8.4	7.0	4.9	5.9
	First order	0.12	1.3	0.9	1.8	1.5	0.4	1.0	0.1	0.7	0.3
	Hermite	0.31	1.2	0.7	0.2	1.6	0.6	0.8	0.0	0.0	0.1
	Spline	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Harmonic times		11	12	13	14	15	16	17	18	19	20
Relative error (%)	Zero	0.9	0.9	1.9	7.8	5.6	4.0	3.9	5.6	2.5	14.5
	First order	1.7	1.4	1.4	2.5	1.8	2.0	1.9	2.7	3.1	3.9
	Hermite	0.0	0.1	0.1	0.7	1.0	0.8	0.4	0.5	0.2	0.4
	Spline	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

tion accuracy of the cubic spline interpolation is several orders higher than other methods.

It can be seen that the zero interpolation and the first-order linear interpolation have huge harmonic calculation errors around the fundamental wave. This is because fundamental amplitude is very different from the harmonics, which harms the interpolation results. As the harmonic frequency increases, the errors will increase, too. Accuracy is not stable. The piecewise cubic hermite interpolation is one among cubic interpolations, mainly based on its first derivative. Its accuracy is higher than that of the above two methods, i.e., the overall error is within 2%, but not stable enough to meet the requirements of the power system. The cubic spline interpolation has more limitation than the hermite, and it makes the accuracy of the spline higher. Its average relative error is 3.31E-5. It is suitable for the accuracy requirements of the power system. From the above, the cubic spline interpolation is selected as AFT harmonic analysis interpolation instead of the previous method.

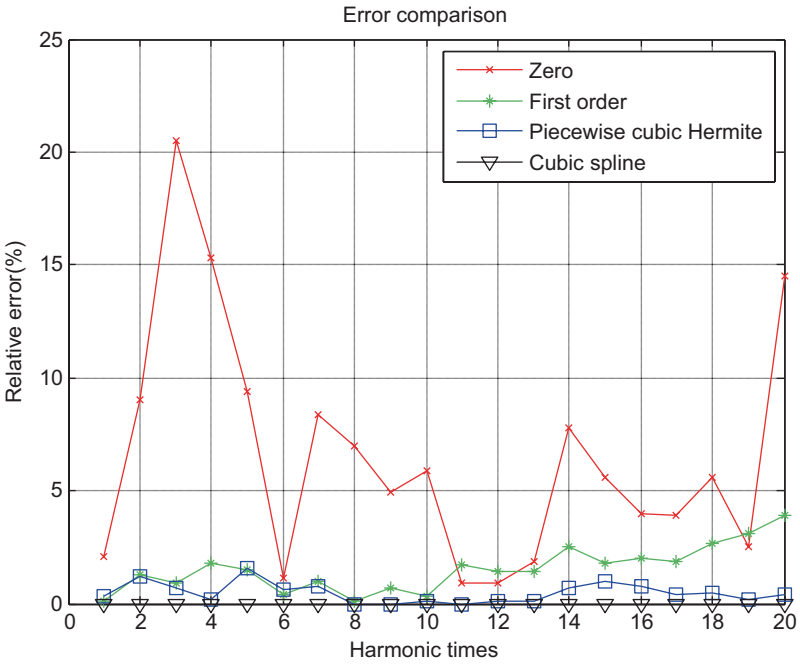


Fig. 29.5 Errors' comparison

29.5 Conclusion

In the power system, the number of sampling points is fixed to 200, which is not an integer power of 2. That does not make FFT so suitable for harmonic analysis. This chapter proposes a method to analyze harmonics in the power system by AFT. AFT has advantages such as simpler calculation, better parallelism, higher accuracy, and no limitation on the number of the sampling points. However, its application is limited by the interpolation. This chapter uses four interpolation methods for AFT harmonic analysis. The result is that the cubic spline interpolation has least influence on the accuracy of AFT harmonic analysis, achieves the requirements of the power system, reduces the limitation of AFT, and provides new ideas of power system harmonic analysis.

References

1. Xue H, Yang R. A novel method for non-integer harmonics measurement using continuous wavelet transform. Autom Electr Power Syst. 2003;27(5):49–53. (In Chinese).
2. Pang H, Li D, Zu Y, Wang Z. An improved for harmonic analysis of power system using FFT technique. Proc CSEE. 2003;23(6):50–4. (In Chinese).

3. Eren L, Devaney MJ. Calculation of power system harmonic via wavelet packet decomposition in real time metering. Instrumentation and Measurement Technology Conference, Proceedings of the 19th IEEE. USA; 2002. p. 1643–47. (Vol. 2)
4. Huang W, Dai Y, Quan H. Harmonic estimation method based on Daubechies wavelet. Trans China Electrotech Soc. 2006;21(6):45–8. (In Chinese).
5. Zhang X, Wu J, Jiang Z, Chen G. An algorithm for computing DFT using arithmetic Fourier transform. ACTA Electr Sin. 2000;28(5):1–3. (In Chinese).
6. Li T, Cheng S, Yang M. Power system harmonic analysis based on Hilbert–Huang transform. Proc CSEE. 2008;28(4):109–13. (In Chinese).
7. Zhang X, Xu Y, Chen G. Practical implementation of the arithmetic Fourier transform. J Electr Inf Technol. 2004;26(6):935–9. (In Chinese).
8. Irving S, Reed M-S, Truong TK, Hendon E, Tufs DW. A VLSI architecture for simplified arithmetic Fourier transform algorithm. IEEE Trans Signal Process. 1990;40(5):1122–32.
9. Zhang X. Piecewise cubic Hermite interpolation function and its application to hydrology. South-to-North Water Transf Water Sci Technol. 2009;7(5):92–4. (In Chinese).
10. Liu C. The convergence of the piecewise cubic hermit interpolation. J Xian Inst Technol. 1994;14(2):151–4. (In Chinese).
11. Chen W, Wang Z. The application of cubic splines interpolation in the project fitting. J Cent China Norm Univ (Nat Sci). 2004;38(4):418–22. (In Chinese).

Chapter 30
Distribution System Reliability Assessment Based on the Improved Monte Carlo Method

Xiaofang Sun and Huijia Liu

Abstract In order to assess the reliability of complicated distribution networks, the complicated network structure is simplified and stored in some serial numbers in the computer according to the network topology and the function of the components in the distribution system; at the same time, a new efficient assessment method for distribution network reliability is presented based on the Monte Carlo simulation method and the numbering method in this chapter. The application of the proposed method to the subsystem F4 of RBTS-Bus6 displays its efficiency.

Keywords Power distribution networks · Reliability assessment · Numbering method · Monte Carlo simulation

30.1 Introduction

Power distribution system is directly connected with the customers and located at the end of electrical power system. With the improving of the customer's requirement for the reliability of power supply, it is very important to assess the reliability of distribution networks.

Several methods have been developed to assess the reliability of distribution networks, such as grouping into the analytical method [1–3] and the simulation method [2–7]. The analytical method is based on failure mode and effect analysis (FMEA), and the simulation method can be divided into the non-sequential Monte Carlo simulation method [6] and the sequential Monte Carlo simulation method [7]. The advantage of the analytical method is that these methods have high accuracy and fast speed with the disadvantage of the basic data requirement; besides, it is difficult to model the complex distribution system. As to the advantage of the simula-

X. Sun (✉) · H. Liu
Electrical Engineering & Renewable Energy School, China Three Gorges University, 443002 Yichang, China
e-mail: 956521681@qq.com

H. Liu
e-mail: epallhj@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_30

tion method, it is flexible in application, can avoid the disadvantages of the analytical method and reflect the running state of system in a truer manner. The accuracy of the simulation method is proportional to the length of the simulation time, but the simulation method features the unique advantage of large distribution systems.

As the executive components determine the fault effect range, the breakers, fuses, isolating switches and loop switches are deemed as boundaries to determine the fault effect region and form zones. Thus, distribution systems can be simplified and numbered into some serial numbers, and a new approach for reliability assessment of complicated distribution systems is proposed based on the Monte Carlo simulation method and the numbering method in this chapter.

30.2 Sequential Monte Carlo Method Based on State Duration Sampling of the Component

The state duration sampling of the component is a sequential Monte Carlo method. The failure time of the components and the repair time of the components obey exponential distribution with the equations shown as below:

$$\begin{cases} f(t) = \lambda e^{-\lambda t} \\ g(t) = \mu e^{-\mu t} \end{cases} \tag{30.1}$$

where λ is the failure rate of the component (f/yr); μ is the repair rate of the component (f/yr); $f(t)$ is the failure probability of the component of t moment; $g(t)$ is the repair probability of the component of t moment.

$$\begin{cases} F(t) = 1 - e^{-\lambda t} \\ G(t) = 1 - e^{-\mu t} \end{cases} \tag{30.2}$$

where $F(t)$ is the probability of the component failure time less than t; $G(t)$ is the probability of the component repair time less than t.

$$\begin{cases} F'(t) = 1 - F(t) = e^{-\lambda t} \\ G'(t) = 1 - G(t) = e^{-\mu t} \end{cases} \tag{30.3}$$

where $F'(t)$ is the probability of component working time more than t; $G'(t)$ is the probability of component repair time more than t.

$$\begin{cases} TTF = -\dfrac{1}{\lambda} \ln F'(t) = -\dfrac{1}{\lambda} \ln R(t) \\ TTR = -\dfrac{1}{\mu} \ln G'(t) = -\dfrac{1}{\mu} \ln R(t) \end{cases} \tag{30.4}$$

where TTF = the failure time of the failed component; TTR = the repair time of the failed component; $F'(t)$ and $G'(t)$ are all random numbers, $R(t) = F'(t) = G'(t)$.

30.3 Realizations of Feeder Partition and Zone Numbering Method

30.3.1 Feeder Partition

A representative distribution network has three main types of executive components: disconnecting switch, fuse, and circuit breaker, which can isolate the fault and restore the power supply of the non-fault zone.

The zone is defined as the maximum combination of feeder line segments, transformer and load, which are connected with each other with the switches as the boundary. Every load point at the same zone has the same interruption duration time and the same failure rate, and the distribution network is formed by zone and executive components.

The complicated distribution network is displayed in Fig. 30.1. The zone is separated by the switching element, and the simplified distribution network of F4 is shown in Fig. 30.2. All feeders, including the distribution transformers and LP18

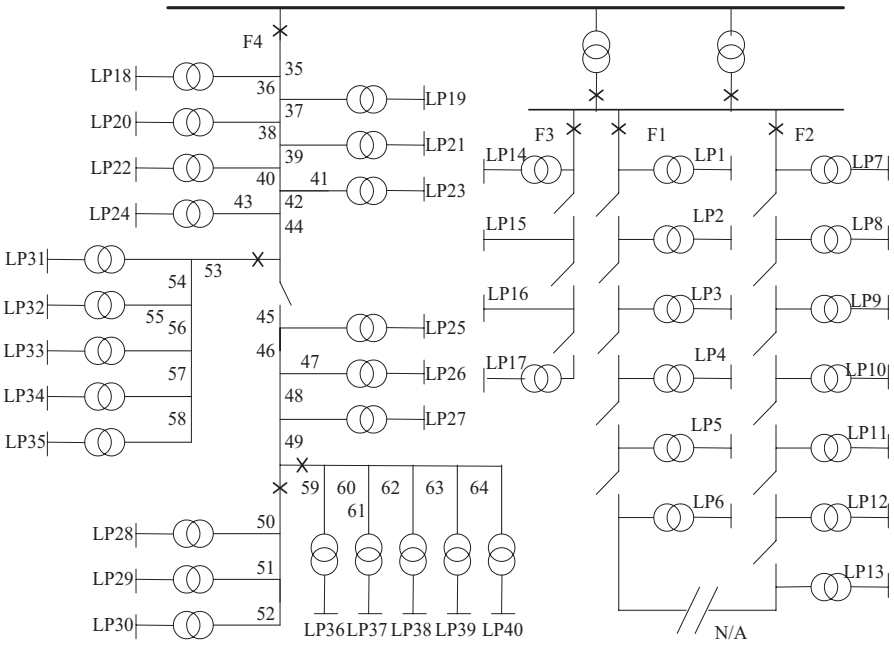


Fig. 30.1 IEEE RBTS-Bus6 test system

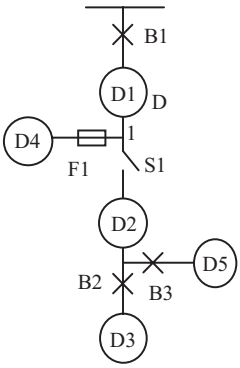


Fig. 30.2 Simplified distribution network of F4

~LP24 as separated by B1, B3, S1, constitute the first zone D1. Zone D1 and zone D2 are separated by S1, while zone D2 and zone D3 are separated by B2.

30.3.2 Numbering Rules of Zones

Based on the feeder partition of distribution networks, all the zones are numbered, and the zone numbering method presented in this chapter can simplify the structure of distribution network, which is easy to be stored and processed by the computer. A new effective evaluation method for distribution system reliability is proposed based on the Monte Carlo simulation method and the zone numbering method.

Firstly, the simplified distribution network is stratified; the network layering method is shown in Fig. 30.3. The next is numbering the zones, and the number of a zone is represented with one-dimensional vector, which is used for describing the topology of the zone; the first bit of the vector represents the feeder interval number. The specific form of each bit of the vector is shown in Fig. 30.4.

The meaning of each bit of the vector is shown below:

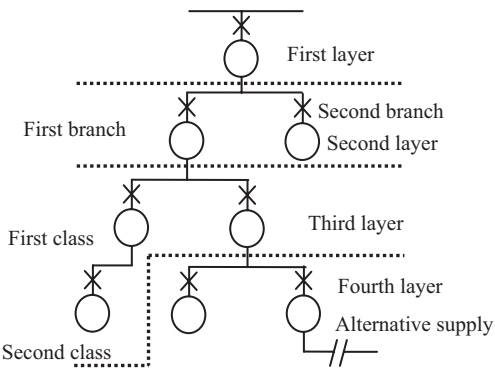


Fig. 30.3 Pre-processing of the simplified distribution network

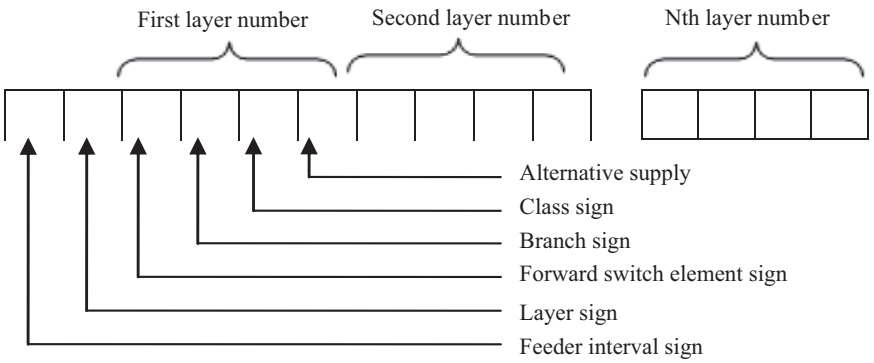


Fig. 30.4 Specific form of each bit of the vector

1. The feeder interval sign indicates the feeder interval number.

2. Layer sign indicates the layer number of the zone.
3. Upward switch element sign indicates the upward switch element of the zone. If this bit is "1," it represents the disconnecting switch; if this bit is "2," it represents the circuit breaker.
4. Branch sign indicates the zone branch number.
5. Class sign indicates the zone class number, representing which class the zone is from.
6. Alternative supply sign: if the zone has alternative supply, the bit is "1"; otherwise, the bit is "0."

30.3.3 Fault Effect Zone

The radial distribution system is divided into the following four different regions as shown in Fig. 30.5. The fault effect can be determined by checking the bits of the zone vector.

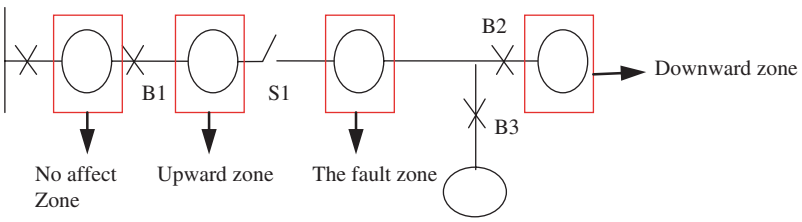


Fig. 30.5 Fault effect zone

1. The fault zone: The fault zone contains the fault component; the interruption frequency of the load points in this zone increases by 1, and the interruption duration time is the repair time of the failure component.
2. Upward zone: If the upward switch element of the fault zone is the disconnecting switch, the zone between this disconnecting switch and the operating breaker is defined as the upward zone; the interruption frequency of the load points in this zone increases by 1, and the interruption duration time is the isolating time of the fault zone.
3. Downward zone: The downward zone to the fault zone is defined as the downward zone. The interruption frequency of the load in this zone increases by 1, and the interruption duration time is determined by whether the zone has the alternative supply or not.
4. No effect zone: The zone not affected by the fault is defined as the no effect zone.

30.4 Realization of Algorithm Based on Zone Numbering Method and Monte Carlo Method

In combination with the Monte Carlo method and the zone numbering method, a new method of reliability evaluation is developed. In the Monte Carlo simulation process, the failure time and the failure sequence order of the various components are simulated. Based on the zone numbering method, the fault element of the zone is searched and the position of the zone is located, and the fault consequences are determined; then, the consequences of failure records are returned to the Monte Carlo simulation program. Finally, upon the simulation calculation for a period of time, the reliability indexes are produced. The steps of the method are shown below:

Step 1: Input the original data.
Step 2: Simplify the network and number the zones.
Step 3: Generate n random number between 0 and 1 and calculate the corresponding TTF by equation (4).
Step 4: Search the minimum TTF and calculate the corresponding TTR by Equation (4).
Step 5: Use the numbering method and determine the fault effect region.
Step 6: Change the time sequence, make $TTF_i = TTF_i + TTR_i + TTF_i'$ and form the new time sequence with no fault components.
Step 7: Check the simulation time; if the time limit does not exceed, return to step 3; otherwise, stop the simulation.
Step 8: Calculate the reliability indexes of the load point and system.

According to the zone numbering method and the Monte Carlo method, the reliability indexes of load points and the distribution system can be calculated.

30.5 Example Study

The IEEE RBTS is a representative electrical distribution system, and its subsystem F4 demonstrated in Fig. 30.1 is selected as the test object in this chapter. This subsystem has a main feeder, two lateral feeders and 30 load points. The corresponding data of all the components and load points are provided in [8]. The repair time and failure rate of the feeder line are 5 h/f and 0.05 f/yr km, respectively; the failure rate and repair time of the transformer is 0.015 f/yr and 20 h, respectively; the switch operation time is 1 h. Assume that all the switch components such as fuse, breaker, and isolating switch can take correct action.

Apply the algorithm presented in this chapter to the example system. The set time is 5000 years, and the calculation result of the reliability assessment of the subsystem F4 is shown in Tables 30.1 and 30.2.

Compared with the analytical method [9] as shown in Tables 30.1 and 30.2, the method proposed in this chapter shows that the deviation of load point reliability indexes is 0.01 %~4.81 % and the deviation of the system reliability indexes is 0.01 %~3.26 %. The maximum deviation is less than 5 %. The calculation results demonstrate the efficiency and accuracy of the approach presented in this chapter.

30.6 Conclusion

According to the function of the switching element, the distribution network can be simplified and the simplified network consists of zones and switch elements. As every component in the same zone has the same fault interruption time, every load

Table 30.1 Load points reliability indexes in F4

Load point	This method			Analytical method		
	λ (f/yr)	γ (h/f)	U (h/yr)	λ (f/yr)	γ (h/f)	U (h/yr)
LP18	1.2648	4.9236	6.2274	1.2900	4.7015	6.065
LP 25	1.2648	6.2587	7.9161	1.3200	6.2272	8.22
LP 28	1.6726	5.8098	9.7175	1.7150	5.7142	9.8
LP 31	1.9556	4.5246	8.8484	1.9550	4.4627	8.725
LP 36	1.9488	5.3230	10.373	1.9724	5.4903	10.83

Table 30.2 Reliability indexes of subsystem F4

Index	This method	Analytical method
SAIFI(f/syst.cust)	1.5975	1.547
SAIDI(hr/stst.cust)	8.1592	7.926
CAIDI(hr/cust)	5.1075	5.002
ASAI	0.9991	0.9991

point in the same zone has the same reliability index. The simplification can largely condense the number of components and cut calculation complexity of reliability assessment. Based on numbering method proposed in this chapter, it is easy to determine the fault effect region. With the zone component as the calculation object, the calculation result of the method proposed herein shows its validity and accuracy, and it can be applied to the reliability assessment of the complicated distribution networks.

References

1. Guo YJ. Power system reliability principle and application [M]. Beijing: Tsinghua University Press; 1986. pp. 169–6 (in Chinese).
2. Billinton R, Wang P. Distribution system reliability cost/worth analysis using analytical and sequential simulation techniques [J]. IEEE Trans Power Syst. 1998;13(14):1245–50.
3. Leite da Silva AM, Schmitt WF, Cassula AM, Sacramento CE. Analytical and Monte Carlo approaches to evaluate probability distributions of interruption duration [J]. IEEE Trans Power Syst. 2005;20(3):1341–8.
4. Li F. A fast approach of Monte Carlo simulation based on linear contribution factors to distribution reliability indices [C]. Transmission and Distribution Conference Expo. IEEE; 2003. pp. 973–7.
5. Leite da Silva AM, Cassula AM, Nascimento LC, Guimaraes ACR. Chronological Monte Carlo-based Assessment of Distribution System Reliability [C]. In Proc. Int. Conf. Probabilistic Methods Applied to Power Systems. IEEE; 2006. pp. 1–7.
6. Geol L, Billinton R. Monte Carlo simulation applied to distribution feeder reliability evolution [J]. Power Syst Res. 1994;29(3):193–02.
7. Wang P, Billinton R. Time-sequential simulation technique for rural distribution system reliability cost/worth evaluation including wind generation as alternative supply [J]. IEE Proceedings-Generation, Transmission and Distribution. 2001;148(21):355–60.
8. Billinton R, Jonnavithula S. A test system for teaching overall power system reliability assessment [J]. IEEE Trans On Power Syst. 1996;11(4):1670–6.
9. Jing A. Reliability evaluation of distribution system considering distributed generation, master thesis [D]. Beijing: North China Electric Power University; 2012 (in Chinese).

Chapter 31
A Computing Tool for Composite Power System Reliability Evaluation Based on Monte Carlo Simulation and Parallel Processing

Tran Ky Phuc, Vu Toan Thang, Pham Hai Binh and Le Thi Thanh Ha

Abstract This paper describes some of our efforts to reduce the execution time of a computing tool, namely *VietReli,* for composite power system reliability evaluation based on sequential Monte Carlo simulation, especially when alternating current (AC) load flow and optimal power flow techniques are used. The approach is an adaptation of the sequential computing algorithm to parallel and distributed environment based on a cluster of multithread and multicore personal computers (PCs) as hardware and Message Passing Application Programming Interface (MPAPI) as software framework. Test results obtained on several actual power systems showed the high performance of the new computing tool in terms of speedup and simulation time. It is also shown that each thread of PC should be responsible for one simulation year at a time and that the number of threads involved on each PC must be defined beforehand for optimized utilization of computing resources.

Keywords Composite power system · Reliability evaluation · Monte Carlo simulation · Parallel and distributed processing · MPAPI

31.1 Introduction

The electric power sector in Vietnam has been undergoing radical changes in its market structures and regulatory laws. In addition, growing demand for electric energy has caused operating conditions close to the physical limits of system

T. K. Phuc (✉)
Vietnam Institute of Energy, Ton That Tung 6, 10000 Hanoi, Vietnam
e-mail: trankyp@hotmail.com

V. T. Thang
Electricity Regulatory Authority of Vietnam, 10000 Hanoi, Vietnam

P. H. Binh · L. T. T. Ha
Hanoi Technology University, Dai Co Viet 1, 10000 Hanoi, Vietnam

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_31

components. A computational tool that allows evaluation of the power system reliability becomes a necessity.

The reliability evaluation can be divided into four levels (Billinton & Allan, 1994; Tran Ky Phuc, et al., 2011)[1, 6]: generation (Hierarchical Level I, i.e. HL-I), generation and transmission (HL-II), generation + transmission + distribution (HL-III) and generation + transmission + distribution + distributed generation (HL-IV). Composite generation and transmission reliability evaluation (HL-II) is the level that draws more interest from the Vietnam industry.

For power system reliability evaluation, as well as analytical approaches, the Sequential Monte Carlo Simulation (SMCS) is one of the most important methods due to its advantages (Billinton & Allan, 1994)[1], such as: more accurate modelling of power system components and operating conditions, providing probability distributions of variables of interest and handling the chronological aspects of system operation. SMCS is especially suitable for systems which are dominantly hydro, like in Vietnam. Literature survey has shown that the reliability evaluation based on SMCS has undergone a great deal of development in the last some decades (Borges, et al., 2001; Lingfeng Wang & Singh, 2009)[3, 7]. To reduce computation time, approximate load flow techniques such as direct current (DC) model are often used in most works. In the case of the Vietnam power system, AC load flow and OPF techniques must be used to obtain more accurate estimates of the line power flows, the bus voltages and the generating unit reactive power limits. Unfortunately however, the simulation approach implemented in sequential computing environment has often caused unacceptable computation requirements.

Substantial improvements are more likely to come from intelligent system methods or parallel processing that is feasible because of naturally decoupled characteristics of MCS (Lingfeng Wang & Singh, 2009) [7]. It is observed that in many works, a parallel version of a MCS algorithm was implemented on parallel computer architectures with distributed or shared memory (Gubbala & Singh, 1995)[4]. Relatively few works have reported on the implementation of parallel computing in distributed systems (Borges, et al., 2001)[3]. However, with the availability of high-speed PC a reliability evaluation system can benefit from this form of decentralized computer architecture due to its geographically distributed nature, flexibility, scalability, cost advantages and great computing speed.

This paper describes our work on *VietReli*—an adaptation of the SMCS sequential computing algorithm for composite reliability evaluation to parallel and distributed processing based on a cluster of PCs as hardware and MPAPI, a message passing implementation(Thomsen, 2008)[9], as the software framework. Performance of this computing tool implemented in a cluster environment will be shown from tests on the IEEE Reliability Test System (RTS) (IEEE Committee Report, 1979) [5] and the Vietnam power system.

31.2 Composite Reliability Evaluation Implemented in Sequential Computing Environment

The sequential Monte Carlo simulation technique is based on sampling the probability distributions of the system component state durations (Billinton & Allan, 1994) [1]. The SMCS based algorithm for composite reliability evaluation can be summarized in the following steps (Billinton & Wangdee W, 2006)[2]:

Step 1 Specify the initial state of each component.

Step 2 For each simulation year sample the system state that consists of load level, generation status and line status.

Step 3 For each hour of the simulation year conduct power system analysis, including: calculate power flow; check if any operating constraint is violated; if yes, corrective action to alleviate constraint violations must be conducted using an optimal power flow (OPF) solution.

Step 4 If the OPF solution shows that there is no load curtailment, go to step 3 for the next simulation hour. Otherwise, accumulate the reliability indices functions $\Phi(S)$ for the simulation year and go to step 5.

Step 5 Repeat steps 3–4 till the end of the simulation year. Calculate the yearly reliability indices functions and update the expected values of the indices $E(\Phi)$ and the accuracy of SMCS.

The expected value of the indices $E(\Phi)$ can be calculated as follows (M. V. F. Pereira & L. M. V. G. Pinto, 1992)[10]:

$$E(\Phi) = \frac{\sum_{i=1}^{NS}\left(\sum_{j=1}^{n_i(S)} \Phi(S_{j,i})\right)}{NS} \tag{31.1}$$

where $n_i(S)$ number of occurrences of system state S in year i,

$\Phi(S_{j,i})$ indices function corresponding to *jth* occurrence in year i,

NS number of simulation years.

The accuracy of SMCS may be expressed by the coefficient of variation β, which is a measure of the uncertainty around the estimates, and is defined as:

$$\beta = \frac{\sqrt{V(\Phi)/NS}}{E(\Phi)} \tag{31.2}$$

where $V(\Phi)$ is the variance of the estimator.

Step 6 If the accuracy of the estimates of the index function β is less than a predefined error ε, the simulation is terminated. Otherwise, return to step 2.

As was noted before, the key feature of this algorithm is that AC power flow and OPF techniques have been used to get more accurate estimates of the power system parameters. Whenever AC power flow method is not suitable, which is mostly due to ill-conditioned network or multi-failure situations, the software automatically uses the DC method to obtain solution.

31.3 Parallel and Distributed Processing Models for Composite Reliability Evaluation based on Monte Carlo Simulation

Parallel processing can be defined as the utilization of multiple hardware components to perform a computation job. A distributed computing system is a collection of autonomous computers, interconnected by a communication network, that work together to satisfy the information processing needs. A cluster of multicore PCs has been selected as basic hardware for our computing tool. For cluster of multicore PCs, the most suitable software model is the master—slave model with message passing interface (MPI) as communication tool (Thomsen, 2008)[9].

As shown above the sequential algorithm of composite reliability evaluation consists of two kinds of iterative loop: through each simulation year and through each simulation hour. Moreover, the simulation process of each hour (or each year) is independent from the process of other hours (years). Thus, the main simulation job can be partitioned in two ways: (i) one year or (ii) one hour *as distributed processing grain*. However, one hour grain partition may cause an increase in communication time and a difficulty in arranging simulation output. Therefore in *VietReli* each thread of PC should be responsible for one simulation year.

Based on the above selected hardware and software models, distributed computing system architecture has been designed for *VietReli* as shown on Fig. 31.1.

The distributed computing algorithm for composite reliability evaluation can be summarized as follows:

- **Registration**: Master Node and Computing Nodes register to Register Node.
- **Simulation preparation**: Master Node sets up simulation parameters through user interface or database, initiates its own simulation environment, and sends general simulation information to the Computing Nodes. Each Computing Node uses general information received from Master Node and its own specific information to set up simulation environment of the Computing Node.
- **Simulation Process**: Each Computing Node sends a message to the Master Node to inquire a simulation task − Master Node sends information about assigned tasks (which operating years to be simulated), to the Computing Node, − Each Computing Node dispatches simulation tasks to its workers. Each worker carries out simulation for assigned year and returns the result to the Computing Node − Computing Node returns simulation results (i.e. yearly reliability indi-

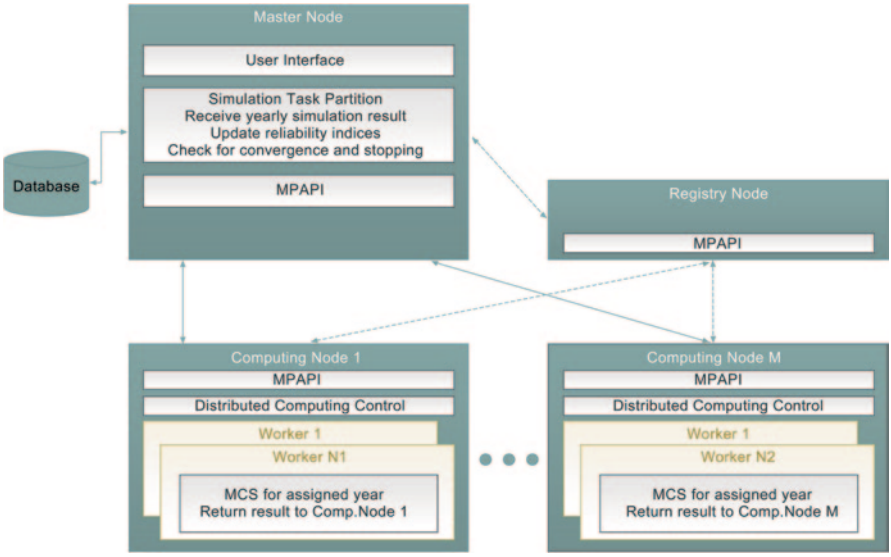


Fig. 31.1 Computing system architecture for *VietReli*

ces) for assigned years to the Master Node − Master Node updates estimate of reliability indices based on results for simulated time span, checks convergence and stopping criterion. This process is repeated until the Computing Nodes receives a termination request from the Master Node.

- **Termination**: Computing Nodes will terminate simulation if they receive a termination request from the Master Node. Simulation results will be presented to users and stored in the database.

31.4 Test Results

The tests are explored in several power systems, mainly the 24 bus IEEE Reliability Test System (IEEE Committee Report, 1979)[5] and VietPS – the Vietnam power system. The communication environment is a Local Area Network with up to 8 PCs, each PC with core i5 processor. The LAN speed is 100 Mbps. The parallel and distributed processing is implemented in Visual C# with TCP/IP protocol for message passing based on a MPAPI framework (Thomsen, 2008)[9].

The testing procedure is described as follows. First, the test systems are examined for reliability evaluation on standalone multicore PC (core i5) using different numbers of threads (from 1–8). Secondly, the test systems are examined on a cluster of 8 PCs (core i5); each PC involves 4 threads for simulation. Performance indices are average computation time per simulated year, speedup and efficiency. Speedup and efficiency are defined by the following equations:

Table 31.1 Simulation time, speedup and efficiency with different number of threads

Test system	Performance indices	Number of threads							
		1	2	3	4	5	6	7	8
IEEE-RTS	Simulation time (s)	193	99	66	48	50	52	57	55
	Speedup (times)	1.00	1.95	2.92	4.04	3.86	3.68	3.41	3.51
	Efficiency (%)	100	97	97	101	77	61	49	44
VietPS	Simulation time (s)	1954	1011	843	725	1066	1724	868	1071
	Speedup (times)	1.00	1.93	2.32	2.70	1.83	1.13	2.25	1.82
	Efficiency (%)	100	97	77	67	37	19	32	23

$$\text{Speedup} = T_1 / T_p \tag{31.3}$$

$$\text{Efficiency}=(\text{Speedup}/p) \tag{31.4}$$

where T_1 and T_p are average simulation time on 1 and p threads, respectively (for test on standalone PC); T_1 and T_p are average simulation time on 1 and p PCs, respectively (for test on cluster of p computers).

These are the observations from the test results:

- Multithreading test on standalone multicore PC: As shown in Table 31.1, for IEEE-RTS, when number of involved threads increases from 1 to 4, the speedup increases from 1 to 4.04, the simulation time is reduced from 193 to 48 s and the efficiency, in general, decreases; when number of threads increases from 4 to 8, the speedup drops and the efficiency decreases at a higher rate. Thus, one should find out the optimal number of threads to optimize utilized computing resource of the PC (in this instance, 4 threads). It is easy to see from Table 31.1 that, the larger system scale is, the lower rate of increase of speedup when number of involved threads increases.
- Cluster test on 8 connected multicore PCs: For IEEE-RTS, when the number of PCs increases from 1 to 8, the speedup increases from 1 to 7.92, the simulation time is reduced from 49.6 to 6.2 s. For VietPS, the speedup increases from 1 to 7.62 and the simulation time is reduced from 729.4 to 95.7 s.

31.5 Conclusion

The execution time of the computing tool (*VietReli*) for actual composite power system reliability evaluation based on sequential Monte Carlo simulation becomes critical and needs to be reduced when AC load flow and OPF techniques are involved in the system state analysis step. It is shown that a cluster of high speed multithread and multicore computers and message passing MPAPI frame-

work are a suitable parallel processing environment for improving the computing tool due to availability and low cost of hardware and simplicity in programming. Analysis has shown that each thread of PC should be responsible for one simulation year at a time. Tests on IEEE-RTS and the Vietnam power system have shown that the computation time is reduced with the increasing number of PCs in cluster, and for each PC the number of threads involved in simulation should be defined beforehand to optimize the utilization of computing resources.

References

1. Billinton R, Allan R. Reliability evaluation of power systems. 2nd ed. New York: Plenum Press; 1994.
2. Billinton R, Wangdee W. Predicting bulk electricity system reliability performance indices using sequential Monte Carlo simulation. IEEE Trans Power Deliv. 2006;21(2):909–917.
3. Borges CLT, Falcao DM, Mello J, Melo A. Composite reliability evaluation by sequential Monte Carlo simulation on parallel and distributed processing environments. IEEE Trans Power Syst. 2001;16(2):203–209.
4. Gubbala N, Singh C. Models and considerations for parallel implementation of Monte Carlo simulation methods for power system reliability evaluation. Power Syst IEEE Trans On. 1995;10(2):779–787.
5. IEEE Committee Report. IEEE reliability test system. IEEE Trans Power Appar Syst. 1979;PAS-98(6):2047–2054.
6. Pereira MVF, Pinto LMVG. A new computational tool for composite reliability evaluation. IEEE Tran Power Syst. 1992;7(1):258–264.
7. Phuc TK, Thang VT, Ha LTT. A Monte Carlo simulation based computing tool for assessing reliability of generation systems. J Sci Technol Tech Univ. 2011;(85):7–12.
8. Thomsen F. Introduction to message passing and the MPAPI framework. 2008. [Online].
9. Wang L, Singh C. Reliability evaluation of composite power systems using parallel genetic algorithms: Some conceptual and simulation studies. PSCE ’09, IEEE/PES; 2009. pp. 1–7.

Chapter 32
Review of Voltage Sag Disturbance Recognition

Hao Yang, Jianwen Li and Yonggang Li

Abstract In this chapter, the characteristics of various types of voltage sag disturbances are analyzed. The domestic and foreign research on voltage sag disturbance recognition methods is reviewed both in process and results. The methods based on direct parameter classification, wavelet decomposition and neural networks, S-transform and similarity classification, support vector machine (SVM), and other classification methods are discussed in detail. Finally, the present problems of voltage sag disturbance recognition methods are analyzed, and future research is discussed.

Keywords Recognition · Review · Voltage sag disturbance

32.1 Introduction

With the growth of power technology and the large-scale application of sensitive equipment, harm caused by voltage sags attracts more and more attention. As one of the most serious failures, it happens when the root mean square (RMS) voltage decreases to a value between 10 and 90 % of nominal voltage for one half-cycle to 1 min [1]. Voltage sags may be caused by many reasons. So the voltage sag sources which can be detected correctly are necessary for the analysis and control of power quality. Currently, the study of the problem has formed relatively mature thoughts: First, characters of disturbance signals are analyzed and extracted by a variety of methods, such as time-domain analysis, frequency-domain analysis, time–frequency-domain analysis, mathematical statistics, and artificial intelligence. Then, the voltage sag sources are classified by variety of techniques, such as artificial neural networks, template similarity matching, support vector machines (SVMs), classification expert system, and fuzzy classification. Including some improved methods, wavelet transform has become an effective tool for feature extraction, especially its singularity detection and multiresolution analysis capability [2]. As the development

H. Yang (✉) · J. Li · Y. Li
School of Electrical and Electronic Engineering, North China Electric Power University, 071003 Baoding, China
e-mail: xuehao_yang@foxmail.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_32

of the continuous wavelet transform and short-time Fourier transform, S-transform has a good time–frequency-domain analysis capability. The signal amplitude over time and frequency constitutes the S-mode matrix, analysis and calculation of the matrix achieves the feature extraction and quantification [3]. The practical application of artificial neural networks is increasing, which has a strong ability to pattern classification, nonlinear function approximation, and adaptive learning [4]. After a number of sample trainings, voltage sag source can be classified by artificial neural networks. Similarity template matching distinguishes the voltage sag sources based on the similarity and differences of sag signals [5]. The classification principle of SVM can be simply interpreted as follows: The optimal separating hyperplane is constructed by different kinds of samples in mapping high-dimensional space or original space, and then the classification is achieved [6].

32.2 Characteristics of Voltage Sag Disturbance Source

Voltage sags may be caused by short-circuit faults, transformer energizing, and induction motor starting. In this chapter, characteristics of the various types of voltage sag disturbance are analyzed by an example of RMS voltage waveform. Figure 32.1 shows the waveforms of three-phase RMS voltage during four common faults.

In Fig. 32.1, from the view of symmetry of three-phase voltage, the waveforms of three-phase short-circuit fault and induction motor starting are symmetrical, while the waveforms of single-phase short-circuit fault and transformer energizing are asymmetrical. From the view of the trends of three-phase voltage waveform, voltage of fault phase mutates at the starting and ending of the short-circuit fault, and it keeps unchanged essentially during voltage sag. Voltage of induction motor starting suddenly reduces at the starting time of sag and remains unchanged for a

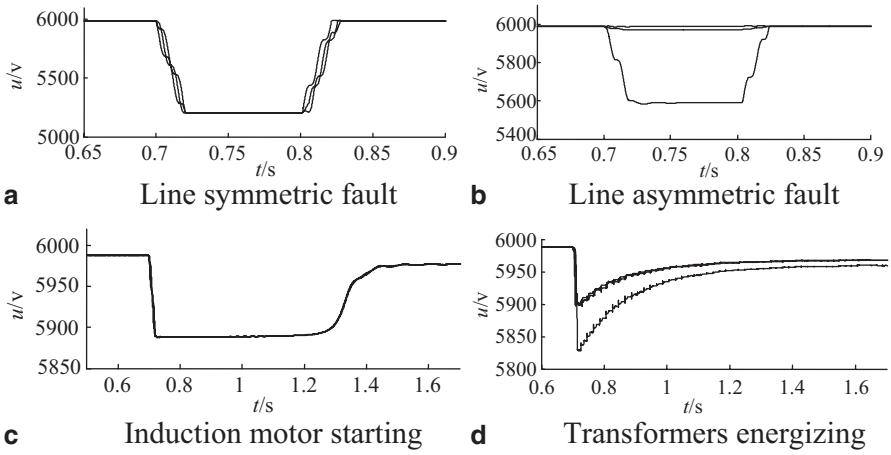


Fig. 32.1 Waveforms of three-phase voltage (RMS) during four common faults [14]

period of time and then rises slowly. Voltage of transformer energizing suddenly reduces at the starting time of sag and then rises slowly. Accordingly, RMS voltage waveforms of different types of voltage sags have the distinct characteristics, and voltage sources can be identified based on these features.

To obtain the characteristics of voltage sag, methods to detect voltage sags such as peak voltage detection [7], RMS detection [8], Fourier transform [9], the traditional Direct Quadrature (DQ) transform [10] and its improved method [11] are required.

32.3 Methods of Voltage Sag Disturbance Recognition

Currently, research methods of voltage sag disturbance recognition can be attributed to the following: First, signal processing or theory and mathematical statistics are adopted to extract features. Secondly, the voltage sag sources are classified based on these features. In this chapter, the methods are divided into five categories: direct parameter classification, wavelet decomposition and neural networks, S-transform and similarity classification, support vector machine (SVM), and other methods.

32.3.1 Direct Parameter Classification

In this method, power parameters are calculated and analyzed directly for the classification. Then, the voltage sag sources are classified based on threshold setting. The process is simple and straightforward and easy to implement.

The RMS voltage during sag is analyzed by the RMS method [12]. Sag depth, beginning and ending time, and any other characteristics are obtained. The effectiveness of the full-cycle RMS algorithm and the half-cycle RMS algorithm is analyzed, and then the delay error curve of the start time and end time of voltage sags is obtained. And a more accurate voltage sag duration could be obtained after the error correction by the curve. The voltage sag source is identified by the computational analysis of RMS voltage [8, 13]. As shown in Fig. 32.1, the RMS voltage waveform contains many characteristics such as sag depth, starting and ending time, recovery characteristics, voltage temporarily rise, duration, and phase voltage balance. These features could be first extracted by calculation and then analyzed comprehensively by some determination process: If sudden changes exist during the voltage recovery process, it can be identified as short-circuit fault. Otherwise, it can be identified as transformer energizing or induction motor starting. If voltage swells occur, or the duration of sag is less than three cycles, it can be identified as short-circuit fault. If the three-phase voltage is balanced, it can be identified as three-phase short-circuit fault or induction motor starting; otherwise, it can be identified as asymmetrical short-circuit fault or transformer energizing. If one-phase sag depth is less than 7% of the rating, then it can be identified as short-circuit fault; if the three-phase sag depth is greater than 85%, then it can be identified as transformer energizing or

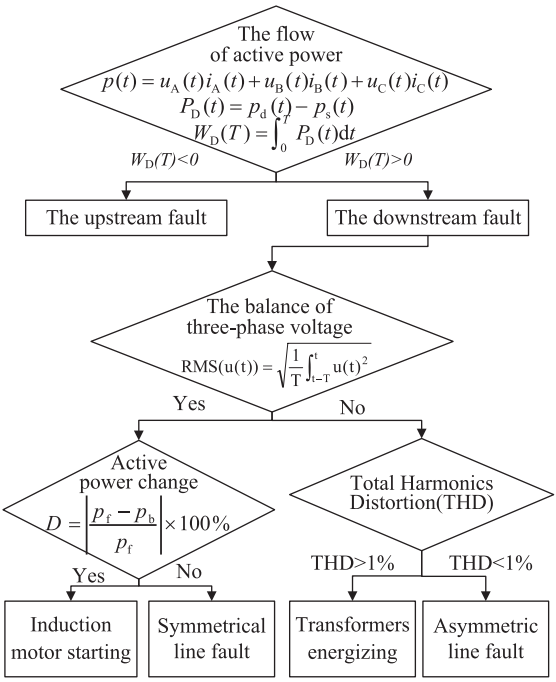


Fig. 32.2 Flowchart of the voltage sag source location and recognition [14]

induction motor starting; if it is between 70 and 85%, then it can be identified as asymmetrical short-circuit faults. The voltage sag sources are identified based on the analysis of the voltage and current at the measurement point [14]. First, based on the current changes of the energy of the system when voltage sag occurs, the position of the voltage sag source relative to the measurement point is determined. Then, the characteristics of different voltage sag sources are analyzed from three-phase voltage balance, sag depth, active power change, total harmonic distortion (THD), etc. Finally, the voltage sag sources are identified based on the judgment process such as that shown in Fig. 32.2.

32.3.2 Wavelet Decomposition and Neural Networks

The methods based on wavelet decomposition and neural networks are widely used [15]. The wavelet transform is a time–frequency-domain analysis methods, which are with good localization properties.

Adopting singularity detection theory and modulus maxima theory of wavelet transform, singular point and its location can be detected [16]. Wavelet entropy is a combination of wavelet transform and information entropy. It has the unique advantages of wavelet transform in dealing with the irregular abnormal signal and the statistical properties of the signal complexity of information entropy [17, 18].

Artificial neural networks are formed by a large network of interconnected processing units. It is a simplification, abstraction, and simulation of the human brain and reflects the basic characteristics of the human brain [19].

The principles of different voltage sags are analyzed, and a method based on wavelet transform is presented [2]. Based on wavelet transform singularity detection principle, waveform characteristics of voltage sags are extracted by B-spline dyadic wavelet transform. If there are two modulus maxima Mt, they can be identified as non-multilevel sags, and if three Mt occurred, they can be identified as multilevel sags. If $|Mt1|>|Mt3|$, upper multilevel sags happened; otherwise, there are lower multilevel sags.

Defined that

$$R_{fd}=\left|\frac{(M_{t1}+M_{t2}+M_{t3})}{\max(M_{t1}\cdot CM_{t2}\cdot CM_{t3})}\right|\le 8\% \tag{32.1}$$

If $R_{fd}\le 8\%$, voltage sags of line fault happened. The waveform similarity R_{sim} is adopted to determine the type of the fault. Comparing the two waveforms, if R_{sim} 20 %, the result was similar; otherwise, it was dissimilar. Similarly, the threshold value of three waveforms is set to 30 %. Finally, the advantages of this method to the direct comparison method are shown by the way of contrast.

The sources of voltage sag are identified by wavelet entropy and probabilistic neural network. The signals which contain wavelet energy spectrum entropy and wavelet coefficient entropy are obtained by the wavelet decomposition. The process of wavelet entropy feature extraction is shown in Fig. 32.3. As shown in Fig. 32.4, 150 samples of three types of fault are inputted into probabilistic neural network for training. And the recognition accuracy is improved by adjusting to the extension parameter S, thus achieving the accurate categories of voltage sag sources. Feature vectors are extracted by DB6 wavelet, and then the wavelet neural network is used to classify the source voltage sag [20]. A recognition method is proposed based on the wavelet neural network for an actual grid [21]. First, the disturbance signal is decomposed by wavelet decomposition, and then the disturbances are classified by the information of signal energy in each band. Finally, the neural network is trained by the BP algorithm for which the learning rate is adaptive, thus achieving classification. A method is proposed based on Hilbert–Huang transform and wavelet packet energy spectrum [22]. After applying mode decomposition to the signal, an intrinsic mode function set is obtained. And then set the n-layer wavelet decomposition and calculate the wavelet energy spectrum; the Hilbert–Huang (HH) spectrum including mutation point, amplitude, harmonics, and other characteristics is obtained. Finally, voltage sag sources are classified by the analysis of its HH graph: If there are two point mutations, it is determined that it is a short-circuit fault, and if every phase gets two point mutations, it is determined that it is a three-phase short-circuit fault; otherwise, it is an asymmetric short-circuit fault. If there are harmonics, it can be identified as transformer energizing. If the spectrum is a straight line, it can be identified as induction motor starting.

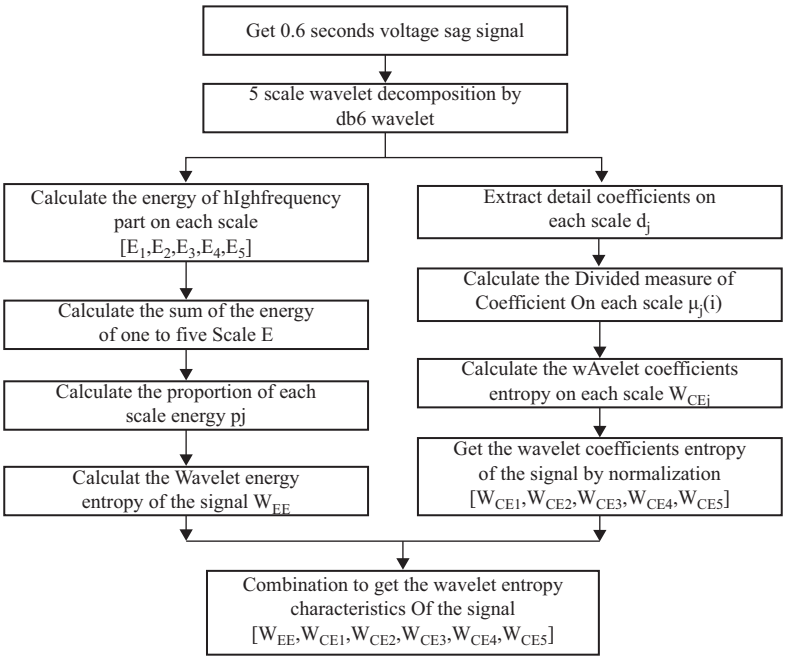


Fig. 32.3 The flowchart of wavelet entropy character extraction [4]

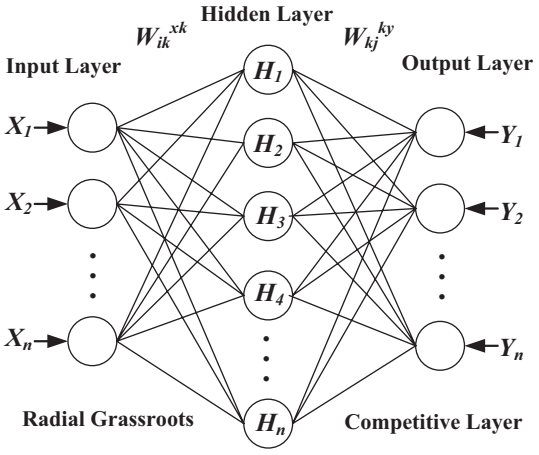


Fig. 32.4 The structure of probability neural network [4]

32.3.3 S-Transform and Similarity Classification

As the development of the continuous wavelet transform and short-time Fourier transform, S-transform was first proposed by Stockwell. It is a reversible part-time-frequency-domain analysis method [23]. Researchers have made some

improvements to the S-transform, by introducing a regulatory factor λ, which can get better time–frequency resolution without increasing the amount of calculation [24].

Template matching similarity is a classification method, which compares the measured signal and standard template established by the known signal, and calculates the similarity, and then determines the categories of the measured signal based on the maximum similarity principle.

The sag signals are analyzed in the time–frequency domain [25]. First, three-phase amplitude convexity, harmonic components, mutation frequency amplitude, phase jump, and other characteristics are extracted. Then, the characteristics are analyzed based on the S-transform time–frequency contour map, first-order central moment B_F of S-mode matrix amplitude envelope vector, and the phase curve of amplitude envelope vector: If the second, third, and fourth harmonics are obviously contained, it can be identified as transformer energizing. The fault phase is identified by B_F, and then the asymmetric short circuit and fault self-clearing are classified. The three-phase short circuit and induction motor starting are classified by the existence of a phase jump. A voltage sag classification expert system is established based on the S-transform. In this system, characteristics of different types of voltage sags are extracted and quantified by the S-transform. And the inference algorithm and discrimination rules of the expert systems are established based on the amplitude factor, harmonic incremental, mutation frequency of amplitude, and the maximum phase increment. A method is proposed based on the improved S-transform, which contrasts the inhibition of the signal noise by the S-transform and improves the S-transform [26]. This method adopts the improved S-transform to analyze and extract the characteristics of the voltage sag sources. The determination rules are formulated based on the binary form parameters including three-phase balance B_1, sag continuing ratio R_1, mutation times N_s, phase jump f_{aij}, even harmonic incremental H_{od}, and sag times N_1. A method is proposed based on the S-transform and similarity classification [5]. First, the bilinear differential scale transform is adopted to build sag global templates. And the most representative portion is extracted to build the only standard templates. Then, the partial template is obtained by the scale transform and localization of the mold time–frequency matrix of the measured signal. The similarity is calculated as follows:

$$S_{AB} = \frac{\sum_{i=1}^{H}\sum_{j=1}^{W} M_A(i, j) \cdot MB(i, j)}{\sqrt{\sum_{i=1}^{H}\sum_{j=1}^{W} M_A^2(i, j) \cdot \sum_{i=1}^{H}\sum_{j=1}^{W} M_B^2(i, j)}} \tag{32.2}$$

where H and W are the number of rows and columns of the matrix, and $M_A(i, j)$ and $M_B(i, j)$ are the elements' amplitude of the i column and j column of sags measured partial template and sags standard template. Then, the categories of the measured signal are determined based on the maximum similarity principle. Table 32.1 shows the comparison of mean similarity between test and standard template.

Table 32.1 Comparison of mean similarity between test and standard template in 40 dB noise [5]

Standard template category						
40 dB		Fault self-clearing	Single-phase short circuit	Fault type changes	Trans-former energizing	Induction motor starting
Test sample category	*Fault self-clearing*	0.9899	0.9599	0.9626	0.9420	0.9207
	Single-phase short circuit	0.9632	0.9972	0.9737	0.9394	0.9440
	Fault type changes	0.9502	0.9761	0.9965	0.885	0.8781
	Trans-former energizing	0.9448	0.9553	0.9431	0.9968	0.9693
	Induction motor starting	0.9394	0.9546	0.9273	0.9715	0.9998

A method is proposed based on the improved S-transform [24]. According to the results of the transform, standard templates are established, and then the progressive similarity and its average value are calculated line by line. Finally, voltage sag sources are classified based on the average value of similarity.

32.3.4 Support Vector Machine

SVM is a method based on statistical theory, which according to the structural risk minimization principle, and with good generalization ability in small samples. Its basic principle can be expressed as follows: Vector input is transformed to a high-dimensional vector space by nonlinear mapping, i.e., by the inner product with the appropriate kernel, then the classification problems are solved in the new space, and the optimal classification plane is obtained [27].

A method is proposed based on a multiclass SVM [28]. A SVM classification method is proposed based on wavelet transform [29]. The sag sources are classified by a method similar to the SVM —the relevance vector machine [30].

32.3.5 Other Methods

In addition to the abovementioned methods, there are some other methods such as label propagation semi-supervised learning [31], voltage space vector [32], space vector parameters and zero sequence voltage [33], Mamdani fuzzy reasoning [34], phase shift method, and radial basis function (RBF) neural network [35].

32.4 Conclusion

Although the studies of identification methods of voltage sag sources have made some achievements, there are some problems to be studied and resolved. First, in the course of the study, the threshold values are set based on the amount of test data; however, due to the test environment, equipment, and other factors, experimental data may not completely simulate the reality. Secondly, in the actual process of power system operation, voltage sags may be caused by the combined effect of a variety of source sag disturbance. Currently, the studies of voltage sag disturbance recognition mostly focus on the sags caused by a single source. However, the classification results of these methods for the voltage sags caused by a variety of faults superimposed are unsatisfactory, and the research on it is not enough. So the improvement of the accuracy and validity of thresholds and the voltage sag mixed disturbance identification are the issues which are worth studying, discussing, and resolving.

References

1. Xiaomei Y. Research on fault classification method of power quality analysis. Xi'an: Xi'an University of Technology; 2005 (In Chinese).
2. Wang K, Song Z, Chen D et al. Interference source identification of voltage sag in distribution system based on wavelet transform. Proc CSEE. 2003;23(6):29–34 (Beijing) (In Chinese).
3. Yang H, Liu S, Xiao X, et al. S-transform-based expert system for classification of voltage dips. Proc CSEE. 2007;27(1):98–104 (Beijing) (In Chinese).
4. Jia Y, He Z, Zhao J. A method to identify voltage sag sources in distribution network based on wavelet entropy and probability neural network. Power Syst Technol. 2009;33(16):63–9 (Beijing) (In Chinese).
5. Li K, Yang H. S-transform based classification for voltage dips according to standard template similarity. Trans China Electrotec Soc. 2010;25(12):139–53 (Beijing) (In Chinese).
6. Song X. Analysis method of power quality disturbances based on wavelet transform and support vector machine. Harbin: Harbin Institute of Technology; 2007 (In Chinese).
7. Lee S, Vu TK, Cha H. A new fast peak detector for single or three phase unsymmetrical voltage sags. IEEE Energy Conversion Congress and Exposition, Atlanta, Georgia, p. 434–440; 2010.
8. Ding N, Cai W, Suo J, et al. Voltage sag disturbance detection based on RMS voltage method. Asia Pacific power and energy engineering Conference, IEEE Conference publications, Wuhan, p. 1–4; 2009.
9. Zhao F, Yang R. Voltage sag disturbance detection based on short time Fourier transform. Proc CSEE. 2007;27(10):28–34 (Beijing) (In Chinese).
10. Liu Y, Huang C, Ou L, et al. Method for unbalanced voltage sags detection based on DQ transform. Proc CSU-EPSA. 2007;19(3):72–6 (Tianjin) (In Chinese).
11. Fan Z, Liu X. A novel universal grid voltage sag detection algorithm. Power engineering and automation Conference, IEEE Conference publications, Wuhan, p. 1–4; 2012.
12. Wang X, Zhou Y, Liu J, et al. Detection method of voltage sags characterizations using RMS algorithm. Low Volt Appar. 2010;10:48–51 (Shanghai) (In Chinese).

13. Ding N, Cai W, Suo J, et al. Research on voltage sag sources recognition method. Power Syst Technol. 2008;S2:55–9 (Beijing) (In Chinese).
14. Yang H, Li Y, Li J, et al. Location and recognition of the voltage sag source based on the 10 kv distribution system. Electr Meas Instrum. 2014;51(20):64–9 (Harbin) (In Chinese).
15. He W, Yang H. Disturbance classification based on second generation wavelet transform and discrete hidden markov models. Trans China Electrotec Soc. 2007;22(5):146–52 (Beijing) (In Chinese).
16. Lou S. The identification and location of the voltage sag source in distribution system. Beijing: North China Electric Power University; 2007 (In Chinese).
17. Li D, Liu Z, Cai J, et al. Transmission lines fault recognition method based on multi-wavelet packet coefficient entropy and artificial neural network. Power Syst Technol. 2008;32(24):65–9 (Beijing) (In Chinese).
18. Yang J, Luo G, He Z. High voltage transmission line fault classification based on entropy weight of wavelet and support vector machines. Power Syst Technol. 2007;31(23):22–6 (Beijing) (In Chinese).
19. Wei R. Analysis and research on the method to detect voltage sags and recognize its disturbance sources. Changsha: Hunan University; 2012 (In Chinese).
20. Zheng G, Yan XM, Li HW, et al. Classification of voltage sag based on wavelet transform and wavelet network. Machine learning and cybernetics, IEEE Conference publications, Shanghai, vol. 1, p. 466–470; 2004.
21. Chen W, Hao X, Lin J. Identification of voltage sag sources in distribution network using wavelet transform and artificial neural network. J Gansu Sci. 2007;19(4):104–7 (Gansu) (In Chinese).
22. Qi B, Zou J, Fan Y, et al. Sag source identification based on hilbert-huang transform and wavelet packet energy spectrum. Electr Power. 2013;46(8):112–7 (Beijing) (In Chinese).
23. Zhan Y, Cheng H, Ding Y, et al. S-transform-based classification of power quality disturbance signals by support vector machines. Proc CSEE. 2005;25(4):51–6 (Beijing) (In Chinese).
24. Xu F, Liu Y, Yang H, et al. Voltage sags identification based on generalized S-transform. Power Syst Prot Control. 2013;41(17):18–24 (Henan) (In Chinese).
25. Yang H, Liu S, Xiao X. Classification of voltage dips using time-frequency contour and amplitude envelope vectors by S-transform. Trans China Electrotec Soc. 2006;21(8):75–9 (Beijing) (In Chinese).
26. Quan H, Dai Y. Detection and classification of voltage sags based on modified S-transform. J Hunan Univ (Natural Sciences). 2011;38(3):45–50 (In Chinese).
27. Zhang X. Introduction to statistical learning theory and support vector machines. Acta Automatica Sinica. 2000;26(1):32–42 (Beijing) (In Chinese).
28. Lv G, Fang Q, Cai X. Identification of voltage sags source based on multi-class support vector machine. Power Syst Prot Control. 2010;38(22):151–5 (In Chinese).
29. Chen W, Hao X, Lin J. Identification of voltage sags in distribution system using wavelet transform and SVM. Control Autom. 2007;19(4):1605–9 (Beijing) (In Chinese).
30. Lv G, Wu C, Sun W, et al. Detection and identification of voltage sags based on RVM and ANF. Power Syst Prot Control. 2011;39(12):65–8 (In Chinese).
31. Wang S, Lv G. Voltage sag sources identification based on label propagation semi-supervised learning. Proc CSU-EPSA. 2013;25(4):34–8 (Tiajin) (In Chinese).
32. Li B, Wang Y, Xiao X. Voltage sag sources identification method based on voltage space vector. Sichuan Electr Pow Technol. 2012;35(3):14–9 (Sichuan) (In Chinese).
33. Ignatova V, Granjon P, Bacha S, et al. Classification and characterization of three phase voltage dips by space vector methodology. Future power systems, IEEE Conference publications, Amsterdam, p. 1–6; 2005.

34. Li G, Ding N, Xu Y. Voltage sag disturbance recognition based on Mamdani fuzzy inference. J North China Electr Pow Univ. 2010;37(2):43–8 (Beijing) (In Chinese).
35. Lv G, Wang X. Voltage sags detection and identification based on phase-shift and RBF neural network. Shanghai, fuzzy systems and knowledge discovery, 2007(FSKD 2007). Fourth International Conference on Volume 1, IEEE Conference publications, p. 684–688; 2007.

Chapter 33
Automatic Generation of Power System Wiring Diagram Using XSL Based on SVG

Chao Shi, Qianjin Liu and Huiming Xu

Abstract The wiring diagram of a power system is an important guarantee of the system's stable operation and efficient schedule. A novel approach in generating the wiring diagram of the power system in scalable vector graphs (SVG) format by using extensible stylesheet language (XSL) is studied in this chapter, which uses the substation configuration description language (SCL) as the source data. The features and advantages about SVG and XSL technologies are briefly introduced. Then, the necessity and method of constructing the system's graphic element library is discussed. Combined with XSL, the programming process is emphatically expounded. The calculation method of graphic element coordinates and connection relations is eminently suggested. The result demonstrates that the method of using XSL technology to generate the wiring diagram of a power system in SVG format is reasonable and effective. The approach may facilitate the automatic generation of the wiring diagram of a power system and make transformation from SCL to SVG easier.

Keywords Power system wiring diagram · Automatic generation · SVG · XSL · IEC61850 SCL

33.1 Introduction

With the promulgation of IEC61850 and IEC61970 standards, the development and construction of smart grid has been deepened increasingly. XML plays an extremely important role in the exchange of a wide variety of data on the Web and elsewhere because of its distinguishing features such as compatibility and extensibility [1]. It

C. Shi (✉) · Q. Liu · H. Xu
School of Electric Power, South China University of Technology, 510641 Guangzhou, China
e-mail: shichao_0706@163.com

Q. Liu
e-mail: qjliu@scut.edu.cn

H. Xu
e-mail: 56435830@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_33

also highlights huge advantages in data exchange and mining of the power system. With IEC61850 and IEC61970 that employ XML as data carrier [2–3], multiple applications such as data communication, heterogeneous operation, and graphical display can be realized through XML data mining and exploitation. IEC61850 specifies a file format named substation configuration description language (SCL) for describing communication related to IED configurations and IED parameters, communication system configurations, switchyard structures, and the relations between them, which could be used to model and describe the main equipments and topology of the power system [2].

The wiring diagram of the power system expresses the connection relations of equipments attached to various voltage levels and network structure graphically. It is indispensable to the systems' stable operation and efficient schedule; however, the generation and transmission of its wiring diagram are faced with many challenges. On the one hand, the existing wiring diagrams are usually stored in proprietary formats rather than a versatile or a standard format, which has caused a severe problem that the wiring diagrams cannot be reused and it thus results in a waste of resources; on the other hand, the wiring diagrams contain huge amounts of information of the power system, including constant-changing device characteristics and complex topology; meanwhile, it should support dynamic refresh and interactions.

SVG is a language for describing two-dimensional graphics in XML. It supports animation and interaction. SVG has been applied in geographical wiring diagram, power flow display, node voltage display, and fault information system, etc., in the field of power systems. SVG has become an important technical cornerstone in smart grid visualization, and it will be easy to obtain, read, modify, and display the graphical data with the SVG technology. This chapter makes use of IEC61850 SCL as the source data and utilizes SVG and XSL technologies to propose a method of generating the wiring diagram of the power system in SVG format and realize the automatic generation in accord with the proposed method.

33.2 SVG/XSL Overview

SVG and XSL are two key technologies in the automatic generation of the wiring diagram of the power system in this chapter. Both SVG and XSL will be introduced briefly in this section.

33.2.1 SVG

SVG is a language describing two-dimensional graphics and drawing program in XML. It was published by W3C, primarily for the network applications. SVG is composed of three types of graphic objects, namely vector graphics, images, and

text. SVG is especially suitable for describing the power system graphical data. The main advantages are listed below [4].

1. SVG is a text-based vector graphic. It can be created and edited with any text editor, and it can be searched, indexed, scripted, and compressed. SVG is scalable and zoomable and with high quality at any resolution. More conveniently, it can be viewed with the help of IE browser installed with Adobe SVG Viewer plug-in.
2. SVG is purely based on XML; therefore, it inherits the commonality and extensibility of XML. SVG supports animation and interaction, responses to data changes quickly, and refreshes graphic display dynamically. It can provide human–machine interface and realize the simulations of event and fault.
3. SVG is a language describing two-dimensional graphic and drawing program in XML, and it complies with W3C standards (e.g., DOM and XSLT). This makes SVG to have characteristics of saving storage and strong readability.
4. Summarily speaking, SVG meets the demands of the wiring diagrams of the power system. EPRI Control Center Application Programming Interface (CCAPI) recommends SVG as the basic format of graphic interaction in dispatching automation system [5]. The research and application on power system visualization based on SVG is being deepened.

33.2.2 XSL

XSL (extensible stylesheet language) refers to a family of languages used to transform and render XML documents [6]. As XML is designed to transmit and store data without involving specific forms, users can apply XSL to extract data as needed and customize the display forms.

XSLT and XPath are two important parts of XSL family. XSLT (XSL transformation) is a format style description language for transforming XML, while XPath is an addressing language used by XSLT to access or refer to parts of an XML document. XPath specifies the data to be processed. It mainly focuses on node *Substation* and extracts data about device information and their connection relations according to the sequence, test, and loop structures; XSLT provides an objective display style and implements the transformation from XML to SVG [7–8].

33.3 Overall Designs

To generate the wiring diagram of the power system, the initial task is to construct the graphic element library, which can provide graphic styles about different electrical equipments. Then, the main work is programming to process SCL source data. The first step is to import SCL source data and SVG power system graphic element library. The second step is to determine the coordinate of each voltage level and attached equipments by calculation, form a correct and orderly layout of graphic

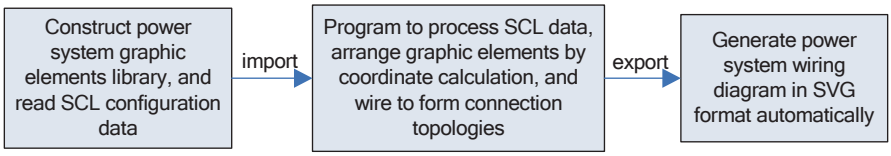


Fig. 33.1 Implementation procedure of wiring diagrams of the power system based on SVG

elements. Finally, an SVG document containing the data about the power system wiring diagram will be exported. Fig. 33.1 shows the implementation procedure.

SCL describes the substation automation system and the relationship between a switch field and the configuration of the intelligent electronic equipment [2]. The node *Substation* in SCL contains the substation structure, the main equipments, and their electrical connection relations as well as the function of corresponding nodes. As shown in Fig. 33.1, SCL data are selected as the source data, and it fully satisfies the requirements of drawing the wiring diagram of the power system, which will be transmitted and stored in the text data instead of proprietary format. It will contribute a lot to the standardization and unification of graphic data processing and display.

33.4 Construction of Power System Graphic Element Library

The power system consists of generator, transformer, transmission line, and a variety of electrical equipments. Such equipments are connected together according to certain rules and unified as a whole. The electrical equipment is an important part of the power system. The wiring diagram contains same type of graphic elements. As the construction of graphic element library covers types of electrical equipments as comprehensive as possible, it is the first work to implement the wiring diagram generation of the power system. Proper definitions and references of graphic elements will reduce the information to be rewritten and the amount of data so that it could make graphic element reuse to achieve a maximum level, reduce the application's calculation amount and the size of graphic files, realize the transmission time and memory consumption reduction, and thus improve the efficiency of the entire system.

A class of graphic element is defined by SVG, and it often groups line, circle, and other basic geometric shapes together in a specific layer order and combination relations. The next task is referring to the SVG schema and electrical equipments' graphical representation specifications and defining graphic elements one by one to construct a graphic element library containing transformer, breaker, disconnector, current transformer, voltage transformer, bus, connecting line, etc.

< g > is a container element for grouping related SVG elements together, and it can contain other < g > elements nested within it to an arbitrary depth [4]. When

there are graphic elements with the same characteristics, they can be abstracted into a same class, defined in < g > and identified with the attribute "*id.*" Then, we can refer and instantiate the graphic elements according to their *id* after definition. The graphic element circuit breaker *CBreak* is defined as follows:

```
<g id="CBreak">
<rect width="12" x="3" height="28" y="4"
style="stroke:rgb(255,0,0);fill:rgb(0,255,0);stroke-width:2;"/>
<circle r="2" cx="9" cy="4" style="stroke:rgb(2555,0,0);fill:rgb(0,255,0);stroke-
width:1;"/>
<circle r="2" cx="9" cy="32" style="stroke:rgb(2555,0,0);fill:rgb(0,255,0);stroke-
width:1;"/>
</g>
```

Configure the element *rect* and *circle*, assign a value to each attribute, and specify the size, position, width, color, and other necessary attributes; then, a graphic element breaker *CBreak* will be constructed. Graphic elements are defined successively, and then a power system graphic element library will be formed finally. When the parts of wiring diagram of the power system are constructed completely, we can assemble the parts together to constitute a complete wiring diagram of the power system.

33.5 Implementation

In this section, the concrete details will be discussed. Firstly, the guideline of implementation method using XALAN processor is illustrated; secondly, the mapping relations and procedure of programming are analyzed. Last but not the least, a typical example is given to testify the correctness of proposed method.

33.5.1 Design of Implementation Method

The languages and tools supporting SVG generation and processing are very rich. C++, Java, Ajax, JSP, and other common programming languages are used to provide an interface; Xalan, Batik, and other toolkits also support SVG.

Xalan is an XSLT processor for transforming XML documents into HTML, text, or other XML document types (including SVG). It implements XSLT and XPath and can be used from the command line, in an applet or a servlet, or as a module in other programs [9]. The method is briefly described in Fig. 33.2.

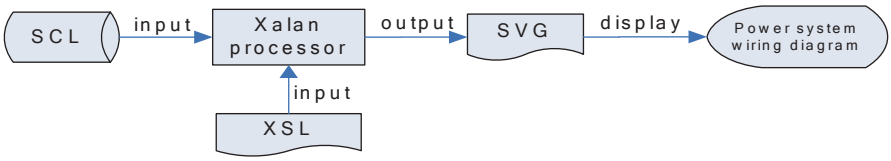


Fig. 33.2 Method of generating wiring diagram of the power system based on SVG by using SCL + XSL

Transformation from XML into SVG could be achieved according to the method shown in Fig. 33.2. An SVG file containing wiring diagram data of the power system will be put out and easily viewed with the help of IE browser.

33.5.2 Procedure

Upon the completion of the construction of graphic element library, the challenge and difficulty of wiring diagram generation is converted to the calculation of coordinate and connection relations. XSL references graphic element according to electrical equipment, determines initial coordinates according to *VoltageLevel* and *Bay*, and draws bus and connecting line according to the element *Connection* and its attributes. The most simple and convenient way to generate SVG file automatically is to enter MS-DOS interface, change directory to which contains XML and XSL files, and run command "java -jar xalan.jar -in exam.xml -xsl exam.xsl -out exam.svg." After the competition of the work listed above, a wiring diagram of the power system with complete equipments, rational layout, and normal connection relations will appear.

SCL contains information about the power equipment and the connection relations. Elements and attributes correspond to the wiring diagram information, respectively. Making full use of the mapping relationships between SCL nodes and the power system wiring diagram will help a lot to determine the graphic element class and the initial coordinate and their connection relations. The mapping relations are illustrated in Fig. 33.3.

XSL programming is the key to the automatic generation of the wiring diagram of the power system. The coordinates and connection relations of graphic elements

```
<Substation>      Select substations interested to generate wiring diagram
  <VoltageLevel  Nam=' '  Voltage=' ' >Draw bus and determine initial coordinate of Bay by VoltageLevel
      <Bay  Nam=' ' >                Consider Bay as a device container, determine x-coordinate of Device
          <Device  Nam=' '  Type=' ' >     Select corresponding graphic element from library by Type
              <Connection  CNodeNam=' ' />     Determine connections by CNodeNam
          </Device>
      </Bay>
  </VoltageLevel >
</Substation>
```

Fig. 33.3 Mapping relations between the SCL nodes and the wiring diagrams of the power system

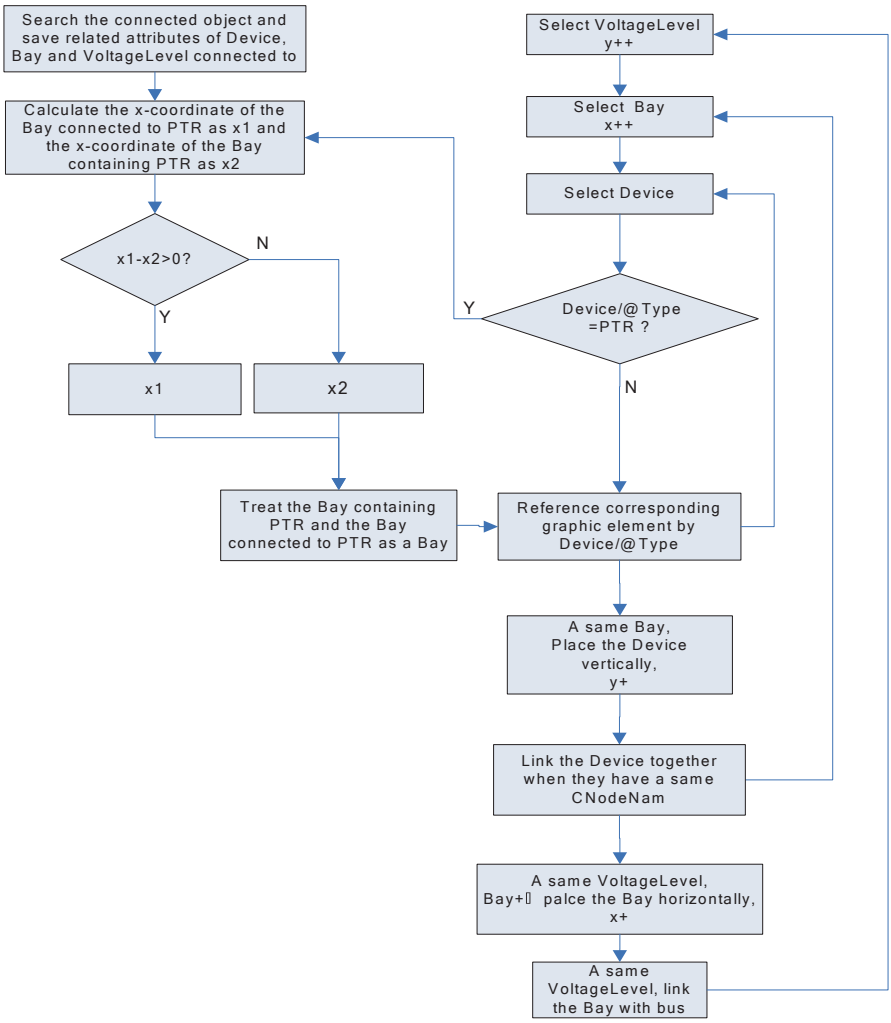


Fig. 33.4 Procedure of generating the wiring diagram of the power system by using XSL based on SVG

are given by XSL programs. Hence, the algorithm to calculate coordinates and connection relations directly affects the accuracy and rationality of wiring diagrams. As SCL data have a tree structure, XSL should drill down by hierarchical relationships when processing SCL data. XPath is employed to address the elements needed. The processing procedure is shown in Fig. 33.4.

The XSL elements used to design algorithm are summarized as follows [8]:

1. Select mode

The XSL element < xsl:for-each > can be used to select every XML element of a specified node set through looping. It is used at a high frequency to select the

element of the same parent node. For example, it is used to traverse all *Device* elements under *Bay*.

The < xsl:value-of > element can be used to extract the value of an XML element and add it to the output stream of the transformation. The device names and voltage levels are extracted by using < xsl:value-of > .

2. Test mode

The element < xsl:if > is used to put a conditional test against the content of the XML file. As used in device judgment, for example, it can be used to test whether the device type is transformer or not.

< xsl:if > provides single conditional test, while < xsl:choose > is used in conjunction with < xsl:when > and < xsl:otherwise > to express multiple conditional tests. The bus color is decided by < xsl:choose > .

3. Match model

An XSL style sheet consists of one or more sets of rules called templates. A template contains rules for implementation when a specified node is matched. XPath is used to address nodes. < xsl:template > defines matching and processing rules. In this chapter, < xsl:template > defines template; < xsl:apply-template > applies a template rule to the current element or the current element's child nodes. < xsl:call-template > calls a named template containing parameters.

33.5.3 Application

The appendix of IEC61850–6 gives out a complete and typical SCD (substation configuration description) file, as configured in SCL, including modules such as *Head*, *Substation*, *Communication*, *IED*, and *DataTypeTemplates*. The configured station is composed of two voltage levels, double buses, and four bays. There are transformers, circuit breakers, disconnectors, voltage transformers, current transformers, and other common power equipments. *Substation* describes the primary system which contains the equipment and the topology information. Extract *Substation* module and process the data by following the instructions in Fig. 33.4. Then, an SVG document containing the contents of the wiring diagram of the power system will be generated automatically. The SVG document can be viewed on IE browser in which Adobe SVG Viewer plug-in is installed. The result is shown in Fig. 33.5.

The result shows that the SVG document generated by XSL is displayed as wiring diagram of the station which configured in SCL. The wiring diagrams display complete equipment with rational layout and correct connection relations. In addition, the display forms can be adjusted according to the users' customization. It can be customized to hide or show the voltage level, bay name and device identifier and other information. The proposed method of generating the wiring diagram of the power system automatically based on SVG and XSL technologies is reasonable and effective.

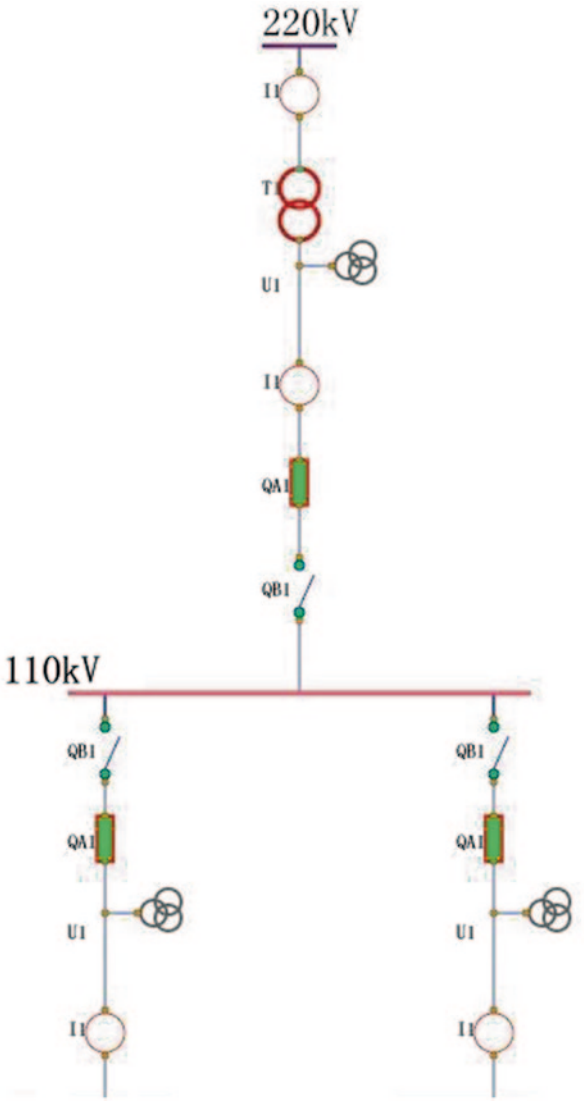


Fig. 33.5 Automatic generation of the wiring diagram of the power system by using SVG and XSL

33.6 Conclusion

The wiring diagram of the power system is an important guarantee of its stable operation and efficient schedule. Application of SVG and XSL provides a novel approach in solving the wiring diagram generation problem. This chapter proposes a method of generating the wiring diagram of the power system in SVG format by using XSL with the main work summarized as below.

1. It employs SCL configuration data as source data to generate the wiring diagram of the power system. SCL, SVG, and XSL are combined to achieve the goal.
2. It constructs graphic elements of the power system by using SVG technology and forms the graphic element library of the equipment as commonly used.
3. It presents the processing procedure and calculation method of coordinate and connection relations.
4. It testifies the correctness and rationality through a typical example. The result demonstrates that proposed method is very competent to solve the wiring diagram generation problem.

SCL contains huge amounts of data about electrical equipments and their connection relations, and it can be used as source data. SVG possesses excellent display performance and can be used as object file format. XSL provides the transformation styles and the processing program. Expansion of the idea and method proposed in this chapter will realize the automatic generation of wiring diagram from SCL.

References

1. Extensible Markup Language (XML) 1.0. W3C recommendation 2008. http://www.w3.org/TR/xml/.
2. IEC 61850. Communication networks and systems in substations—Part 6: configuration description language for communication in electrical substations related to IEDs. 2003.
3. IEC61970. Energy management system applications program interface (EMS-API). 2007.
4. Scalable Vector Graphics (SVG)1.1. W3C recommendation 2011. http://www.w3.org/TR/SVG/.
5. Common Graphic Exchange RFP V2. CCAPI task force 2011. http://cimug.ucaiug.org/default.aspx.
6. The Extensible Stylesheet Language Family(XSL). W3C recommendation 2003. http://www.w3.org/Style/XSL/.
7. XML Path Language (XPath) 1.0. W3C recommendation 2003. http://www.w3.org/TR/xpath/.
8. XSL Transformations (XSLT) 1.0. W3C recommendation 1999. http://www.w3.org/TR/xpath/.
9. The Apache Xalan Project. Apache software foundation 2011. http://xalan.apache.org/index.html.

Chapter 34
Robust Control for Electric Servo System Based on Sliding Mode Disturbance Observer

Jianguo Zheng, Xingquan Wang, Zhiqiang Ma and Tianpeng He

Abstract This chapter proposes a sliding mode disturbance observer (SMDOB) to deal with the uncertain factors, including modeling errors, parameter perturbations and external interferences faced by the electric servo system control. Then, a proportional integral derivative (PID) angular position tracker is designed to achieve the function of angular position-tracking performance of electric servo system. Thus, the SMDOB and PID tracker form the robust compound control framework of the electric servo system. The computer simulation results indicate that, compared with the traditional PID control scheme, the control scheme proposed in this chapter can ensure the electric servo system to have higher precision of position tracking and stronger robustness against the unknown and time-varying uncertain dynamics in system.

Keywords Electric servo system control · Robust control · Compound control · Sliding mode control · Disturbance observer

J. Zheng (✉) · X. Wang · Z. Ma
Communication Training Base of the PLA General Staff Headquarters,
075100 XuanHua, China
e-mail: anhuizjg@163.com

T. He
BeiHang University, 100191 Beijing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_34

34.1 Introduction

Direct-current (DC) motors have been widely applied to various angular position servo systems, including missile rudder [1], flight simulator [2] and optical tracking platform [3]. However, due to the complex technological environment, the demands on the performance of the DC motor servo system become higher and higher. However, because of the existence of many uncertainties in the electric servo system, including modeling errors, parameter perturbations and disturbance torques, the further improvement of system tracking performance will be hindered.

Among all the controllers applied in the actual projects, the proportional integral derivative (PID) controller is the most widely used, owing to its simple structure, good stability, reliable operation and adjustment convenience. It has also been widely studied and applied in DC motors [4]. However, PID controller lacks robust property against uncertain dynamics in system; therefore, it needs to explore novel control strategies in order to make the motor servo system stronger.

The closed-loop robust control method based on disturbance observer (DOB) has a strong inhibitory effect on the external interferences and a small range of parameter variation, which can significantly improve the performance of control system. Thus, it has become a basic control method for the high-precision servo system and is widely used in various types of electric servo systems [5–6]. Because the essence of the sliding mode variable structure control (SMVSC) is nonlinear control, it has a strong robustness against parameter perturbations and external interferences [7–9]. Therefore, this chapter introduces SMVSC into the construction of disturbance observer, namely the sliding mode disturbance observer (SMDOB), in order to cope with the uncertain dynamics existing in the electric servo system. As a result, it enhances the robust property of the control system, and then the angular position-tracking performance of servo control system can be greatly improved.

In order to synchronously guarantee the engineering practicability and method effectiveness (related to high tracking precision and strong robustness), we intend to explore a new servo control method for electric servo system in this chapter. The control scheme, which is constructed by an inner-loop disturbance observer and an outer-loop PID tracker, can not only deal with the impacts of external disturbances and parameter perturbations but also guarantee the high accuracy of position tracking. All mentioned above is the main starting point of this chapter.

34.2 Problem Description

The control object of the DC motor servo system is the DC motor. For general meaning, the schematic diagram of the DC motor is shown in Fig. 34.1.

The relationship between the counter electromotive force e and the angular speed ω can be described as follows:

$$e = K_e \omega, \tag{34.1}$$

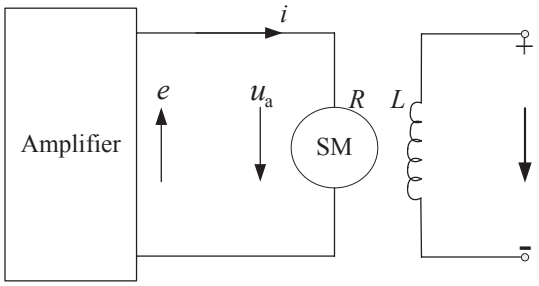


Fig. 34.1 The schematic diagram of the DC motor

where K_e is the voltage feedback coefficient. Moreover, the relationship between the motor torque T and the armature current i can be described as follows:

$$T = K_t i, \tag{34.2}$$

where K_t is the current–torque coefficient.

Consider the effect of the damping torque, and assume the connection between the rotor and the axis is rigid, then the moment equation of the DC motor is as follows:

$$T - B\omega = J\frac{\mathrm{d}\omega}{\mathrm{d}t}, \tag{34.3}$$

where J is the rotation inertia, and B is the mechanical damping parameter.

According to Kirchhoff's voltage law, we can get the following voltage equation:

$$L\frac{\mathrm{d}i}{dt} = u_\mathrm{a} - e - iR, \tag{34.4}$$

where L is the armature inductance, u_a is the input voltage of the DC motor, and R is the armature resistance.

Besides, the relationship between the controller output voltage u and the motor input voltage is as follows:

$$u_\mathrm{a} = K_u u, \tag{34.5}$$

where K_u is the amplification factor of the power amplifier.

The state-space expression of the electric servo object can be obtained by the formula (34.1)–(34.5), and it is shown as follows:

$$\begin{cases} \dot{\theta} = \omega \\ \dot{\omega} = -\dfrac{1}{T_m}\omega + \dfrac{K_m}{T_m}u + d \end{cases}, \tag{34.6}$$

where θ represents the rotation angle of the DC motor; d represents the equivalent disturbances in system, and it is bounded with $|d| < \Delta$; K_m and T_m respectively denote the equivalent magnification factor and the equivalent time constant, and $K_m = K_u/(K_e + BR/K_t)$, $T_m = JR/(K_e K_t + BR)$.

Until now, the state-space model of the electric servo object has been established, and this model provides a theoretical basis for the next design process.

34.3 Robust Compound Scheme Design

34.3.1 SMDOB Design

If we can estimate the equivalent disturbance d in system (34.6) by some means of observation strategy, and then make the same amount of compensation in the control input u, the influences of equivalent disturbances on the control system will be effectively suppressed. Based on this control thought, the next step is to design a SMDOB based on the electric servo object model (34.6), and the designed SMDOB can be regarded as the inner-loop controller.

Introduce variable γ and construct it as follows:

$$\dot{\gamma} = -\frac{1}{T_m}\gamma + \frac{K_m}{T_m}u + y, \tag{34.7}$$

where y is the input signal of SMDOB.

Choose the sliding switching function as:

$$s = \gamma - \omega. \tag{34.8}$$

According to (34.6), (34.7), and (34.8), we can get:

$$\dot{s} = -\frac{1}{T_m}s + (y - d). \tag{34.9}$$

Thus, we can design the SMDOB input as follows:

$$y = \left(\frac{1}{T_m} - k\right)s - q\,\mathrm{sgn}(s), \quad k > 0, \ q > 0. \tag{34.10}$$

It can be proved that when $q > \Delta$, we can get $s\dot{s} \leq 0$, and if and only if $s = 0$, $s = 0$ is tenable. The detailed process of this proof is not given here. Furthermore, $s \to 0$ can also be obtained, namely, $\gamma \to \omega$.

According to (34.9)–(34.10), the output of SMDOB (i.e., the estimated amount of the equivalent disturbances) can be obtained as follows:

$$\hat{d} = -\dot{s} - \frac{1}{T_m}s + y = -\dot{s} - ks - q\,\mathrm{sgn}(s). \tag{34.11}$$

Equation (34.11) will be introduced in the control input of the system as a disturbance compensation term.

In order to avoid the high-frequency chattering caused by switching item, (34.11) can be rewritten in the following form:

$$\hat{d} = -\dot{s} - ks - q\,\mathrm{sat}(s), \tag{34.12}$$

where sat denotes a saturation function [3].

The SMDOB can only achieve the robust stability of the inner-loop system, but not ensure the position-tracking function of the DC motor system. Therefore, an outer-loop controller based on position error regulation is further needed in the next step.

34.3.2 PID Position Tracker Design

To achieve the position-tracking function of the electric servo system, based on the previously designed SMDOB, this section designs a PID controller based on angular position error regulation. The structure diagram is shown in Fig. 34.2:

where θ_d represents the expected angle output of the electric servo system.

Thus, for the electric servo object (34.6), the control law u consists of two parts, and it can be expressed as follows:

$$u = u' - \hat{d}. \tag{34.13}$$

So far, the design of the compound robust controller has been completed.

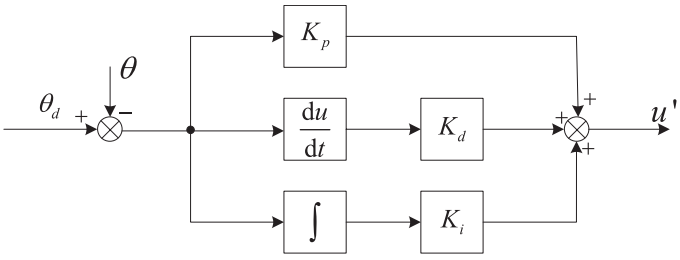


Fig. 34.2 The structure diagram of outer-loop PID controller

34.4 Simulation Results

In this section, we will compare the tracking effects of the proposed control scheme with the pure PID control scheme through computer simulation, in order to obtain the performance advantages of the proposed control scheme.

In the simulation, the actual parameters of the electric servo object are $T_m = 0.025\text{s}$ and $K_m = 150$. However, in the design of controller, the associated nominal parameters are $T_{m0} = 0.0\ 3\text{s}$ and $K_{m0} = 200$.

In addition, the influences of coulomb friction torque and system parameter perturbations are added into the electric servo system. Furthermore, the equivalent disturbances resulted from them are chosen as $d_f = -0.4\,\text{sgn}(\omega)$ and $d_p = 0.1\sin(2\pi t)$.

The coefficients of PID controller are selected as: $K_p = 10.0$, $K_d = 0.2$, $K_i = 0.01$. The coefficients of SMDOB are selected as: $k = 10.0$, $q = 3.0$, $\varphi = 0.1$. At the same time, due to the actual situation, the output of the compound controller is limited between ± 10 V.

In this simulation, the angular position-tracking signal of the electric servo system is shown in Fig. 34.3. Moreover, we compare the tracking errors of the DC motor under the two control schemes, and it is shown in Fig. 34.4.

As seen from Fig. 34.4, compared with the traditional PID control scheme, the position-tracking precision of the DC motor is much better under the proposed control scheme. Concretely, the positive maximum error and the negative maximum error are respectively 0.0758 deg and −0.0756 deg under the traditional PID control scheme. Meanwhile, those are respectively 0.0309 deg and −0.0198 deg under the proposed control scheme, and always maintain at an expected level during the whole control period. This reflects the stronger robust property of inner-loop SMDOB in restricting the equivalent interferences of control system, and then the performance advantage of the proposed control scheme is presented.

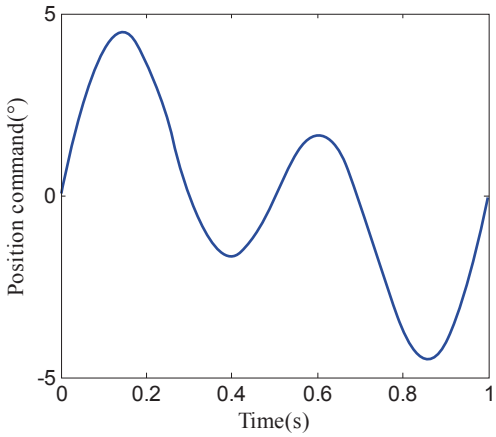


Fig. 34.3 The angular position-tracking signal of the electric servo system

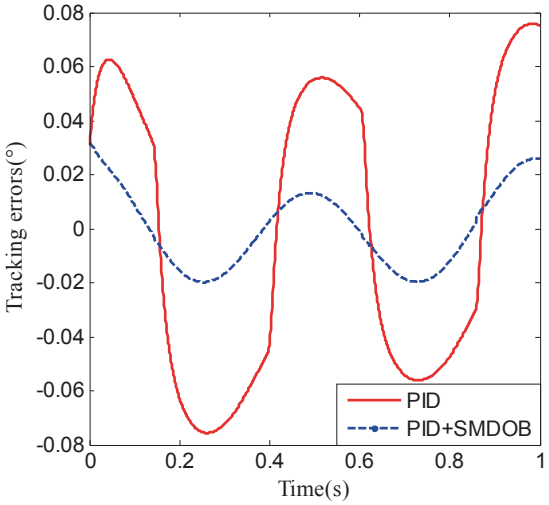


Fig. 34.4 The tracking error curves under the two control schemes

In addition, the control input of the DC motor under the proposed control scheme is shown in Fig. 34.5.

As seen from Fig. 34.5, there is no obvious high-frequency chattering in the control input of the DC motor control system, which can reduce the bad impact on the motor mechanical structure and avoid the instability problem of control system caused by the high-frequency chattering.

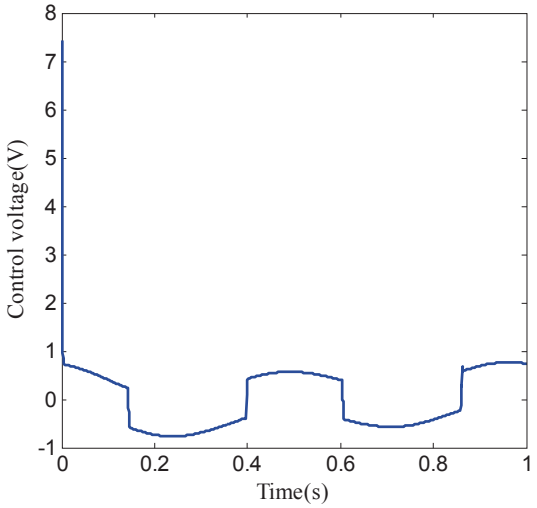


Fig. 34.5 The control input curve of the DC motor under the proposed control scheme

34.5 Conclusion

System uncertainties including modeling errors, parameter perturbations, and external disturbance can seriously affect the performances of electric servo tracking system. Therefore, this chapter presents a robust compound control scheme based on SMDOB. Through computer simulation, we can observe that compared to the conventional PID control scheme, the control scheme proposed in this chapter possesses stronger robustness in response to the unknown and time-varying uncertain dynamics. Moreover, the position-tracking errors of the electric servo system are smaller. In addition, the structure and the algorithm of the proposed control scheme are simple; therefore, it also has practical value in engineering.

References

1. Mianzo L, Peng H. Output feedback H∞ preview control of an electro-mechanical valve actuator. IEEE Trans Control Syst Technol. 2006;1(1):1–9.
2. Wu YJ, Liu XD, Tian DP. Research of compound controller for flight simulator with disturbance observer. Chinese J Aeronaut. 2011;24(5):613–21.
3. Ren Y, Liu ZH, Liu XD, Zhang Y. A chattering free discrete-time global sliding mode controller for opto ronic tracking system. Math Probl Eng. 2013;1:1–8.
4. Liu HS, Chen YC. Simulation and experiment study of BLDCM PID control system. J Syst Simul. 2009;21(16):5157–60 (In Chinese).
5. Lee HS, Tomizuka M. Robust motion controller design for high-accuracy positioning systems. IEEE Trans Ind Electron. 1996;43(1):48–55.
6. Feng Y, Yu XH, Han FL. High-order terminal sliding-mode observer for parameter estimation of a permanent-magnet synchronous motor. IEEE Trans Ind Electron. 2013;60(10):4272–80.
7. Gao WB, Hung JC. Variable structure control of nonlinear systems: a new approach. IEEE Trans Ind Electron. 1993;40(1):45–55.
8. Yu XH. Sliding-mode control with soft computing. A survey. IEEE Trans Ind Electron. 2003;56(9):3275–85.
9. Li JN, Su HY, Zhang YB. Chattering free sliding mode control for uncertain discrete time-delay singular systems. Asian J Control. 2013;15(1):260–9.

Chapter 35
Simulation Analysis of Impacts of Flow Characteristics on Power Oscillation

Jin Tan

Abstract In order to analyze the impact mechanism of the valve flow characteristic on power oscillation, flow characteristics deviation of the steam turbine control valve is firstly analyzed and a coefficient for the deviation assessment is proposed; then, the model of generating unit coupled with dynamic system and electrical system is established according to the characteristics of deviation coefficient, with which the valve flow characteristics deviation is considered. Effect of flow deviation on power oscillation is analyzed by time-domain simulation. The simulation results show that the poor partial flow characteristics of the turbine valves are likely to cause power fluctuations and both the amplitude and the speed of oscillations are positively related to the extent of flow characteristics deviation. The conclusion can provide reference to research on the causes of power oscillations.

Keywords Turbine · Flow characteristics · Governing system · Power oscillation

35.1 Introduction

Previous studies have shown that the turbine's periodic disturbance can easily induce the forced oscillation of the power grid. It is of greater possibility and influence that the forced power oscillations are caused by prime mover disturbance than that caused by load disturbance under the same amplitude and frequency [1–2]. In recent years, although most of the generators in China Southern Power Grid have been installed the power system stabilizer, the oscillations still occur [3].

The valve flow characteristics of steam turbine inlet regulating valve refer to the corresponding relationship between the inlet valve position and the steam flow. If there is a big deviation between the set curve of flow characteristics and the actual valve flow characteristics, there is a risk of power oscillation in operations such as load change and frequency regulation [4–5].

J. Tan (✉)
Guangdong Electric Power Research Institute, 510080 Guangzhou, China
e-mail: tanjin@gddky.csg.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_35

However, little attention has been paid to the influence of flow characteristics deviation on power oscillation. This paper focuses on the effect mode of flow characteristic deviation on power oscillation. The electromechanical coupling model including the flow characteristics is established to find out the key factors of strong correlation which combines the valve flow characteristics and power oscillation.

35.2 Flow Characteristic Deviation

The valve flow characteristics are the correspondence between the valve position and the valve steam flow. The theoretical valve flow characteristics, usually pre-installed in steam turbine control system according to the technical parameters of steam turbine manufacturer, are usually considered corresponding to the actual valve flow characteristics. But there are still deviations between theoretical and actual valve flow characteristics due to valve manufacturing, installation, maintenance, etc. If the differences between the theoretical and the actual valve flow characteristics vary greatly, it can possibly induce load mutation and delays in load and frequency regulation of steam turbine [6, 7].

The flow characteristics deviation might make the regulation system to deliver a deviated valve position demand to regulate valves. Figure 35.1 shows the comparison of the actual and theoretical valve flow characteristics curves of a certain turbo unit; F_t is the theoretical flow characteristic function curve and F_a is the actual flow characteristic function curve.

Observing the difference in quantity of the theoretical and the actual flow curves, the relative deviation ($D_f = (F_t - F_a) \times 100\%$) is calculated. Differences between theoretical and actual valve flow characteristics are represented not only in quantity, but also in rate of change. In different flow reference demand regions, there are different rates of change between the two. In order to compare the difference between the two, the slope ratio coefficient K_f is defined in Eq. 35.1:

$$K_f = \frac{d(F_t)}{d(F_a)}. \tag{35.1}$$

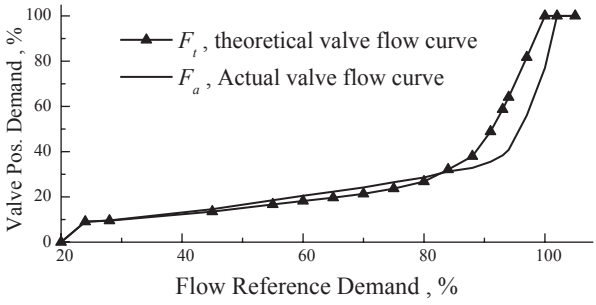


Fig. 35.1 Comparison of actual and theoretical valve flow characteristic curves

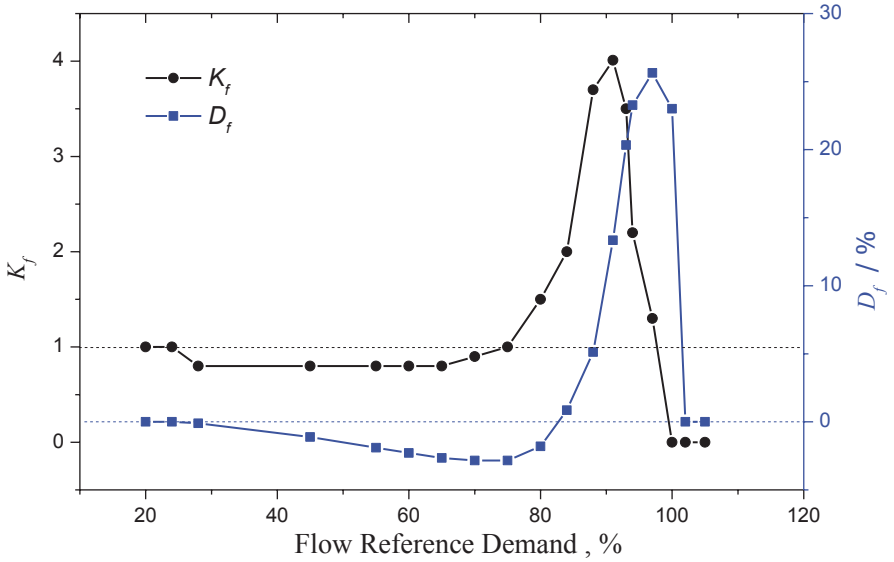


Fig. 35.2 Numeric and slope deviations of set and actual flow characteristics

It is easy to see that $K_f = 1$ and $D_f = 0$, if the theoretical valve flow curve can accurately reflect the actual one.

As shown in Fig. 35.2, when the deviation of the two becomes large, the slope ratio will also become large and can even reach 4 times the actual value. In the flow curve inflection region, both the relative deviation D_f and the slope ratio K_f are obviously large, which implies that the deviation between theoretical and actual flow curves is evidently large in the inflection region.

35.3 Simulation Model

The coupling model between the generating unit and the power system for time-domain simulation analysis is set up by MATLAB, as shown in Fig. 35.3. A single-machine infinite bus (SMIB) system model is adopted for the simulation of generator rotor [8]. With the mechanical power Pm as input, the SMIB system feeds back signals of active power and rotor speed to the digital electrohydraulic (DEH) control system, and then they join into a closed loop. DEH governing system and steam turbine model, respectively, refer to the Institute of Electrical and Electronics Engineers (IEEE) typical models [9–11]. The Valve Position (VP)-flow module in the model is the valve position and flow characteristics relationship of the inlet regulating valve, which is the actual flow characteristics; the flow-VP module is the flow characteristics set in DEH control system.

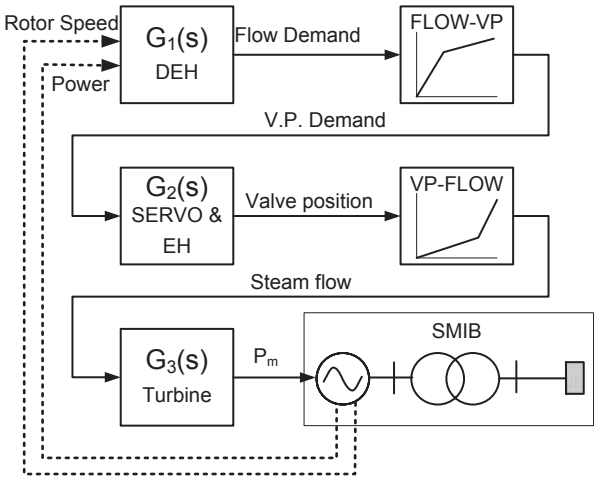


Fig. 35.3 Coupling simulation models considering flow characteristics

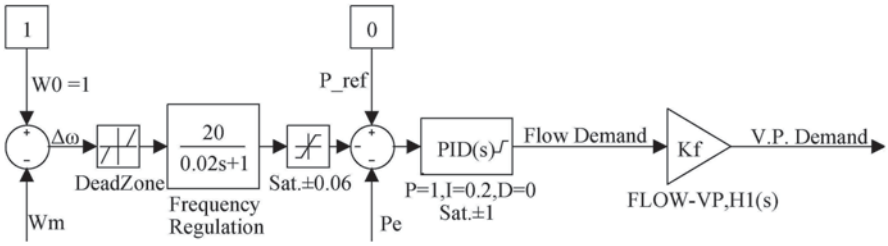


Fig. 35.4 $G_1(s)$, models of DEH control system and flow-VP module

DEH control system puts out flow demand to the flow-VP module and receives the feedback signals of active power Pe and rotor speed Wm from the SMIB system, as shown in Fig. 35.4. For the convenience of simulation, the transfer function of the flow-VP module is set to $H_1(s) = K_f$, as shown in Fig. 35.4. The parameters of the DEH control system are set as below: dead zone, ±0.00067; gain of frequency regulation, 20; time constant of frequency regulation, 0.02 s; output limit of frequency regulation, ±0.06pu; parameters of the power proportional-integral-derivative (PID) controller are $P = 1$, $I = 0.2$, $D = 0$; and output limit, ±1pu.

The signal of valve position demand is converted to the valve position by the electrohydraulic (EH) system (electrohydraulic converter and hydraulic servomotor), and then the valve position is converted to the steam flow by the VP-flow module, as shown in Fig. 35.5. The parameters of the EH model are set as below: the proportional gain of electrohydraulic converter, 2; the integral coefficient of hydraulic servomotor, 1.25; the output limiting, ±1pu; and the time constant of the linear variable differential transformer (LVDT) feedback, 0.02 s. For the convenience of simulation, the transfer function of the VP-flow module is set to $H_2(s) = 1$, as shown in Fig. 35.5.

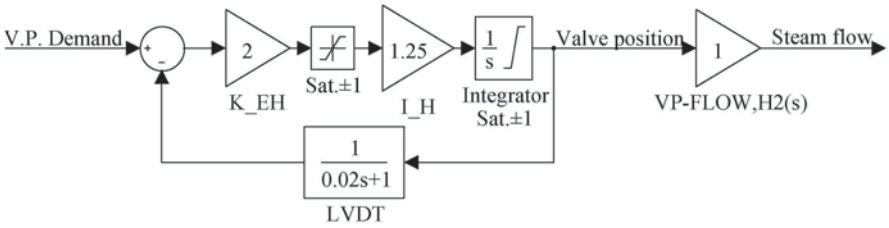


Fig. 35.5 G2(s), models of the electrohydraulic (EH) system and VP-flow module

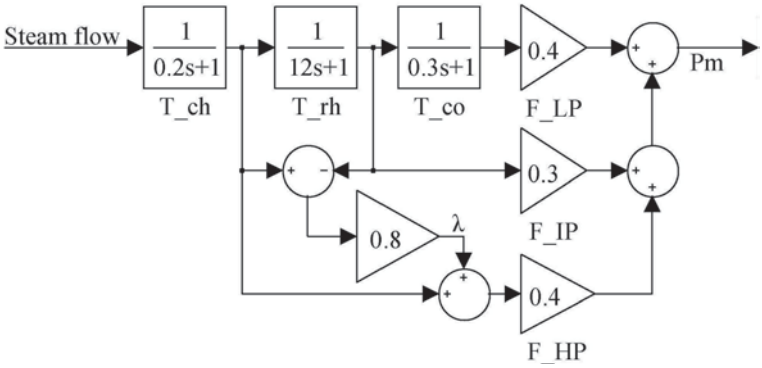


Fig. 35.6 $G_3(s)$, reheat steam turbine model

The model of steam turbine is shown in Fig. 35.6. The parameters of the model are set as below: $Tch = 0.2$; $Trh = 12$; $Tco = 0.3$; $F_{HP} = 0.4$; $F_{IP} = 0.3$; $F_{LP} = 0.4$; $\lambda = 0.8$.

35.4 Simulation Results

The pulse disturbances (with the K_f value set from $K_f = 1$ to $K_f = 2$, 3 or 4) were initiated into the coupling model at the 10th second and removed at the 20th second. The oscillations of rotor speed, power, and load angle were recorded. The simulation results are shown in Figs. 35.7, 35.8, 35.9.

The above simulation results show that the flow characteristics deviation, if it reaches a certain level, is apt to cause low-frequency oscillations of resonance mechanism; in addition, the oscillations will disappear soon after the disturbances of K_f are stopped. With the increase of K_f, the output state changes from stable state ($K_f = 2$) to continuous oscillation ($K_f = 3$, 4); beisdes, the oscillations occur soon and the oscillations' amplitude of rotor speed, power, and load angle is large when K_f is large.

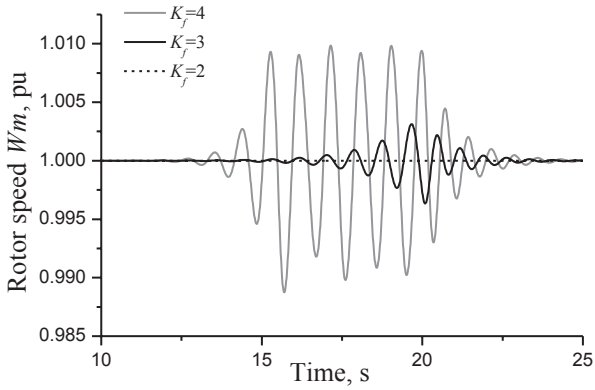


Fig. 35.7 Curves of rotor speed oscillation by pulse disturbance of different K_f values

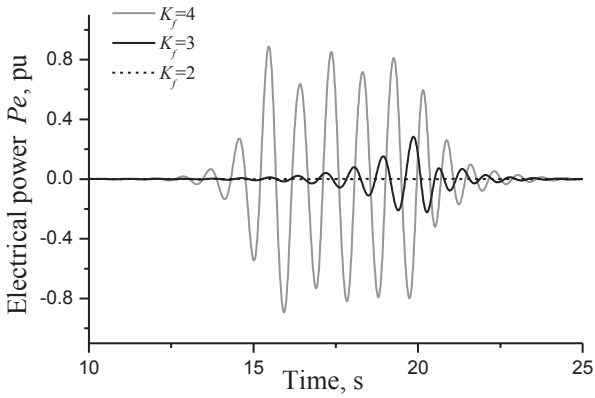


Fig. 35.8 Curves of power oscillation by pulse disturbance of different K_f values

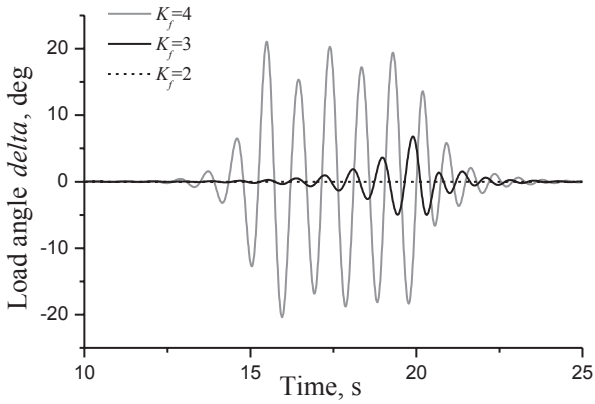


Fig. 35.9 Curves of load angle oscillation by pulse disturbance of different K_f values

35.5 Conclusion

In order to find out the effect mechanism of flow characteristics deviation on low-frequency oscillations, the causes and features of flow characteristics deviation are analyzed, and then the coupling model between thermodynamic system and power system is established for the time-domain simulation analysis. The effects of flow characteristics derivation on the response characteristics of rotor speed, electric power, and load angle are analyzed, respectively, through simulation with the results indicating that the flow characteristics deviation is apt to cause low-frequency oscillations and both the amplitude and the speed of oscillations are related to the extent of flow characteristics deviation. The conclusion can provide reference to research on the causes of power oscillations.

References

1. Tang Y. Fundamental theory of forced power oscillation in power system. Power Syst Technol. 2006;30(10):29–33. (In Chinese).
2. Deng J-x, Hua Y, Zhang F. Study on the nonlinear interactions of oscillation modes. Autom Electr Power Syst. 2003;27(16):35–9. (In Chinese).
3. Xiao M, Liang Z-f. Analysis on the forced oscillation failure in china southern power grid and its handling measures. South Power Syst Technol. 2012;6(2):51–4. (In Chinese).
4. Tian S-f, Shi Z-j, Yan L-t. Study of valve overlapping of steam turbine control system. Turbine Technol. 2008;50(6):448–50. (In Chinese).
5. Li J-b, Liu F-p. Optimization for the flow characteristic function of turbine valves and the influences on safe and economic operation. Electr Power. 2008;41(12):50–3. (In Chinese).
6. Xu Y-h, He R-m, Han Z-y. The cause analysis of turbine power disturbance inducing power system low frequency oscillation of resonance mechanism. Proc CSEE. 2007;27(17):83–7. (In Chinese).
7. Dong C, Yun L, et al. Research on the properties of forced oscillations caused by turbine's periodic disturbance. Power Syst Clean Energy. 2012;28(4):35–41. (In Chinese).
8. Kundur P. Power system stability and control. Beijing: China Electri Power Press; 2002. p. 485–507.
9. IEEE Committee Report. Dynamic models for steam and hydro turbines in power system studies. IEEE Trans Power Appar Syst. 1973;92(6):1904–15.
10. IEEE Working Group on Prime Mover and Energy Supply Models for System Dynamic Performance Studies. Dynamic models for fossil fueled steam units in power system studies. IEEE Trans Power Syst. 1991;6(2):753–61.
11. Tian Y-f, Guo J-y, Liu Y-q, et al. A mathematical model of reheat turbine for power grid stability calculation. Power Syst Technol. 2007;31(5):39–44. (In Chinese).

Chapter 36
Design of Real-Time Tracking System on High-Voltage Direct Current Converter Station Control and Protection System

Wei Hao, Chengjiang Wang, Ming Li and Qingyun Liang

Abstract Diagnostic tool functions of the original fault diagnostic system in direct current (DC) converter stations are not concentrated, and it is unable to accurately diagnose the exceptions of different layers. In order to ensure rapid and accurate location of the host events and the server faults, the real-time diagnosis software for high-voltage direct current (HVDC) converter station control and protection system is designed. The NetDDE technology tracks the suspicious analogy and variables of the switch state using the software interface of MACH2 system, which has been implemented to analyze the exception of the state variable and locate the host fault by recalling the historical data. By calling the stored procedure of SQL server, exceptions can be tracked in a real-time manner by means of the returned data column information. The host and the server exceptions can be rapidly and effectively monitored, tracked, and located by the system. It can also reduce the time of dealing with complex defects and increase the reliability of the whole control and protection system. The system does not affect normal operation of the original system.

Keywords Converter station · Real-time tracking · Diagnostic system · NetDDE technology · Event tracking

36.1 Introduction

The direct current (DC) transmission technology was developed earlier overseas, which was mainly dominated by ABB and SIEMENS. The MACH2 DC control and protection system of ABB was digested and absorbed by some domestic manufacturers (NARI-relays and XuJi) upon introduction into China [1]. It has become the most important technology platform of control and protection system in domestic China, but the diagnostic tools based on this platform have been stuck in the early

W. Hao (✉) · C. Wang · M. Li · Q. Liang
College of Electrical Engineering & New Energy, Three Gorges University,
443002 Yichang, China
e-mail: 178054024@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_36

stage of technology introduction and no company has the secondary development or improvement for it. The fault diagnosis of control and protection system mainly relies on a series of built-in fault diagnosis software of the original SCADA System Engineer workstation, including abnormal packets display tool Intouch, abnormal state monitoring tool Hidraw, DC control and protection system custom wave record tool Inspect, and recorded waveform-viewing tool Reval [2].

Most converter stations run a long time in Hubei Province. These fault diagnostic tools provided by foreign manufacturers have not been updated. The host of Jiangling Converter Station Bipolar Control and Protection system appears fault frequently up to 10 months in 2010. Neither the maintenance staffs nor the manufacturers can locate the abnormal accurately by the diagnostic tools. The function of fault diagnosis is in urgent need of expansion. Specific performances are given below [3]:

1. Unable to diagnose the server fault. DC converter station control and protection system and the database server are tightly coupled and interactive. There is not any tool for the diagnosis to the server so far.
2. Unable to diagnose the host fault. The current diagnostic tools cannot realize the diagnosis function of the abnormal host events, which will lead to disorder of the host event message.
3. Monitoring function of the control and protection system is imperfect. Hidraw only provides a numerical display of state, but it cannot implement continuous recording function of numerical value.
4. The event display interface function is not perfect, and the event display is incomplete. For example, when the event filtering occurs short-term fault, the event information will be ineffective.
5. The original diagnostic tool functions are dispersed, and the interface is unfriendly. Old diagnostic tools are provided by different manufacturers without being integrated. In addition, the interface of original diagnostic tools is in English and simple.

36.2 System Function and Structure Design

36.2.1 Function Design of Abnormal Diagnosis System of Converter Station Control and Protection System

The software system will use the existing software interface of the DC control and protection system by means of manual selection of the continuous monitor, analysis of the suspicious analog and switch state, monitoring the alarm events of the server. It is helpful that the operation and maintenance personnel track and locate all kinds of abnormalities of the converter station.

The system function design is given below:

1. Monitoring and continuous recording in a selective manner of all the state control and protection systems in a real-time manner.
2. Tracking function of the abnormal events of the database server, continuous recording function, and abnormal data modification.
3. Alarm function of limit detection of any state.
4. Historical data monitoring retrieval, browsing, drag, zoom, analysis and derived function.
5. Automatic and manual processing function of historical data.

36.2.2 Logical Structure Design of Auxiliary diagnosis system

The auxiliary diagnosis system of converter can be divided into two parts: the part of host of control and the protection exception diagnosis and the part of database server abnormal diagnosis. According to the functional requirements, the logical structure of the system is designed to have six functional modules: analog monitor module, switch state monitoring module, historical data playback module, historical data disposal module, the SQL event tracker module and database access module. These modules completing the corresponding function operate through the graphical interface of unified diagnostic system.

The system structure is shown in Fig. 36.1.

The whole system adopts a top-down hierarchical modular development process. The whole system is divided into three layers: the user interface layer, the logical function layer, and the data access layer. The user interface provides the user-friendly interface diagnosis system; at the same time, through the user interface, the logical function's layer sub-function modules are called to complete the corresponding abnormal diagnosis function.

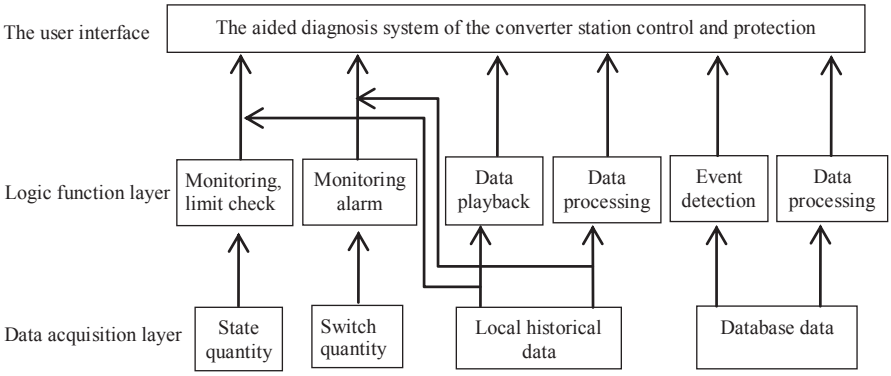


Fig. 36.1 Logical structure of the computer-aided diagnosis system

36.3 The NetDDE Technology

NetDDE is a network dynamic data exchange mechanism. NetDDE technology has been widely used in the industrial control areas to solve the real-time data exchange, such as configuration software (Intouch, RSView32, Fix, Wincc, etc.). The programmable logic controller (PLC) automation system makes the remote connection between PLC and server to realize the data exchange, which is also a kind of application in NetDDE technology; therefore, the application of the converter station data exchange also has high feasibility.

The data acquisition layer provides the data for the logical layer. According to the different data sources, it can be divided into the state acquisition module, database failure information and database data acquisition module.

The host of the control and protection accesses the station within the LAN and sends data to the SCADA through NetDDE protocol [4]. How to use the NetDDE interface to collect data from the host of the control and system is the key to this system. Basic data acquisition process is shown in Fig. 36.2:

In the figure above, the host converts data through NetDDE protocol into data packets that can be spread over Ethernet. In order to achieve the data acquisition process from the diagnostic system to the host, diagnostic system through NetDDE protocol parses the data packet and ultimately converts into data [5]. The key technical difficulty lies in the diagnostic system in NetDDE packet parsing.

There are two main kinds of connection modes in NetDDE: the first is hot link, where server and client data transfer is bidirectional; the second is warn link, where server and client data transmission has a notification process [6]. Warn link will be used for instantaneous fault of swithching value, it will be beneficial to reduce memory usage.

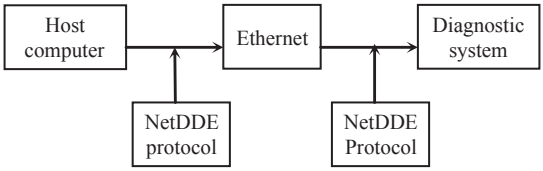


Fig. 36.2 Data acquisition process on the basis of NetDDE

36.4 Functions to Achieve

36.4.1 Monitoring the Running Status

At present, the station switch acquisition is mainly based on Intouch to obtain messages. The Intouch messages can only obtain device fault, but it cannot reflect specific component problem [7]. In addition, aging and mechanical dithering of switching elements may cause instantaneous error action (open or closed). It is difficult for the Hidraw to capture transient fault of the switch.

This module is mainly used to locate the abnormal to lower levels (e.g., card), when monitoring system receives alarm signal from the high level (e.g., electrical cabinet). It is convenient for the maintenance personnel to monitor real-time suspicious state and provide over-limit warning function.

In order to facilitate the operation and maintenance personnel quickly locate the fault, the function module structure is shown in Fig. 36.3.

Monitoring content mainly includes analog quantity and digital switch. On the basis of Intouch alarm message, the operating personnel run query system for monitoring suspicious state.

Monitoring content is divided into the state and switch variables. As to the state variables, the system collects and saves data every millisecond, and draws real-time curve. Data storage is good for the analysis of equipment fault and playback history data. Real-time curve is used to observe the change in trend of state. The curve supports the following functions: pause, continue, enlarge, minify, and capture the dot location of the mouse on curve. The curve coordinate values are mainly used to determine the extent of the amount of the anomaly. As to the switch variables, it only puts out two states of 0 and 1 and collects and preserves a switch status every 1 ms [8].

The state threshold value needs to be set in the limit detection. It will give an over-limit warning and save the data while exceeding the threshold value. The over-limit states are marked in red to prompt the workers, and it uses Access file to save the data.

36.4.2 SQL Event Tracking

The event tracking mainly relies on invoking the stored procedure of system to capture data column information and perform specific function. The tracker is mainly used to adjust and optimize the performance of the database system and monitor the real-time faults. It can optimize CPU, I/O of disk, memory, logical and physical structure of the database, network environment, and SQL statements. It can adjust the partitioning technology of SQL server and rational use of clustering methods. By defining the SQL warning and error events, the data column information of related event is obtained. The server fault tracking is actualized, such as abnormal database connection.

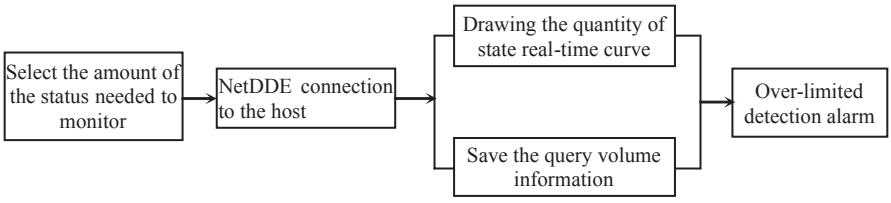


Fig. 36.3 Monitoring logic diagram of the running state

Fig. 36.4 Logical diagram of event tracking

This module is used to monitor abnormal events of the database server [9], such as database connection, communication and warnings or errors in the process of operation data. By tracking the database server warnings and error events, it can avoid no reaction or tucking for the host or server [10]. Because of shortcomings, the converter station itself has produced many warning events. In order to locate the faults, the warning and error events will be filtered by time or user's name [11]. The fault information will be stored in the Microsoft Word file.

In order to locate the server exception conveniently, the module function structure diagram is shown below [12] (Fig. 36.4):

The tracking event module is mainly used to monitor the server exception events. Because the host system itself has some abnormal events, it will produce a series of warning events. In order to locate the abnormal events quickly and accurately, warning and error event data columns (the warning and error events list some relevant information) need to be filtered (e.g., setting up a user name).

It will establish an event tracking module by the SQL statement and get the data columns of the warning and error events by establishing the tracing statements [13] as follows:

sp_trace_create (create a trace definition);
sp_trace_setevent (for adding or deleting a track event column);
sp_trace_setfilter (the filter is applied to the track);
fn_trace_getTab. (trace file information returned in a table format).

The program diagram of the event tracing is shown in Fig. 36.5:

When you start the module, the system will automatically search for all database servers in the form of pull-down menu to the operator for choice. Setting the event data column provides two filters of a user name and time as the user needs. If you do not set the filter criteria, it will monitor the information of all data columns by default. Then, track the event information processing, and Microsoft Word file is used to save the monitored data column information.

36.4.3 Historical Data Processing Module

As the data are collected every 1 ms, monitoring the state quantity for 24 h will produce about 8 GB of date volume. If you do not process the stored data for a long time, it will run out all resources of the hard disk. Hence, it needs to process the stored data regularly; the system data cleaning tips are given every month, and automatically data deletion or compression and removal to another drive package shall be handled by the operator.

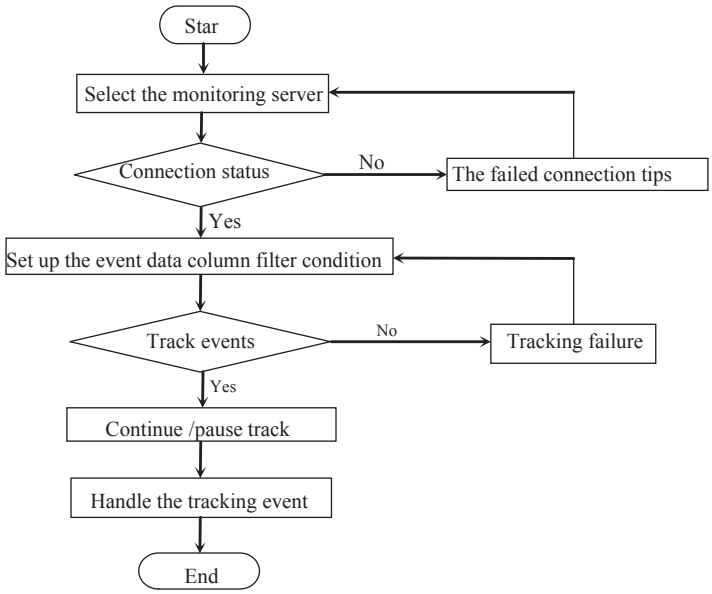


Fig. 36.5 The flow diagram of the event tracing

36.5 Conclusion

With the utilization of NetDDE technology, the system sends data and gets commands to the SCADA of control and protection host station accessing the internal LAN. What is more, it monitors the suspicious state control and protection system of different manufacturers through NetDDE protocol and locates the abnormal levels and equipment or the board quickly and accurately.

Because the database server has been coupled highly with the host system of the protection and control, the system also develops a dedicated tracking module of SQL event. In this way, it can monitor the unusual or abnormal events during the process of data connection, communication, and storage, to avoid the host or server long time no response or jammed fault and other abnormalities for a long time; at the same time, it monitors the operation of the user and abnormal operating event of the server to locate anomalies as soon as possible, which effectively reduces the abnormality investigation time of converter station control and protection system.

References

1. Dai D. Detention method of MACH2 based system of event flow. Hubei Electr Power. 2011;35(6):18–20.
2. Shen Z. Application of MACH2 system in HVDC transmission project. High Volt Eng. 2006;32(9):46–9.

3. LI F. Analysis of control and protection software of high voltage DC transmission system. Shanghai: Shanghai Jiao Tong University; 2006.
4. Tian J. Design and realization of HVDC. Control Prot Syst Electr Power Autom Equip. 2005;25(9):10–4.
5. He H. Application of NetDDE in the automatic network. Microcomput Inf. 2002;6(18):54–7.
6. Zhang J. Method to solve the communication of different configuration software based on NetDDE. Technol Gansu Sci Technol. 2010;26(16):30–337.
7. Wang L. DDE dynamic data exchange technology in measurement experiment. Metrol Measurement Tech. 2009;36(10):19–20.
8. Wang J. The application of power monitoring & controlling system based on InTouch. Tianjin: Tianjin University; 2008. http://www.sqlservercentral.com/stairway/72363/.
9. Li G. Analysis of dynamic modifications to the INI file. Softw Guide. 2011;10(6):26–7.
10. Zhou G. Design and implementation of data in SQL Server database operation tracking system Fujian Computer. Technol Gansu Sci Technol. 2005;23(4):60–2.
11. Sun Q. Effect of SQL profiler in database optimization. Comput Program Skills Maint. 2009;12(4):72–4.
12. Kelly AJ. SQL server profiler or server-side trace. SQL Serv Mag. 2006;14(2):7–8.
13. Stairway to Server-side Tracing—Level 1. http://www.sqlservercentral.com/stairway/72363/.

Chapter 37
Grid Dispatching of Energy Saving and Environmental Protection Based on Thermodynamic System Model with Flow Dimension

Xiao Liang, Jinhui Ma, Duanchao Li and Jianping Wang

Abstract In this chapter, a novel thermodynamic system model with flow dimension is proposed to solve the problem of grid energy saving and environmental protection dispatching (DESEP).The enthalpy entropy flow (EEF) model is used to represent the thermodynamic cycle of the power generator. By adjusting the environmental factors, comprehensive value of energy saving and environmental protection is calculated. It obtains the best output-scheduling scheme by the classical evolutionary programming. Subsequently, the experiments on power generator dispatching are conducted, following three steps: original data analysis, accuracy analysis and efficiency analysis; then, the results show that the values of main factors with coal consumption, boiler efficiency, turbine heat consumption, and auxiliary power ratio under DESEP are close to the best running state.

Keywords Flow dimension · Thermodynamic system · Enthalpy entropy flow model · Grid dispatching · Dispatching of energy saving · Environmental protection

37.1 Introduction

In China, the coal resource is abundant, and the coal-fired generating units accounted for about 68 % of the national total installed capacity. Anhui area is rich in coal and lacks water resource, so the coal-fired units accounted for more than 97 % of the provincial dispatching abilities [1]. In order to realize the policy of national emission control target and reduce the energy consumption and pollutant emissions of

J. Wang (✉) · X. Liang
School of Electrical Engineering and Automation, Hefei University of Technology, 230009 Hefei, China
e-mail: xiaoapollo@163.com

J. Wang
e-mail: wjphfut@126.com

J. Ma · D. Li
Dispatching and Control Center, Anhui Electric Power Corporation, 230022 Hefei, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_37

nitrogen sulfur fumes, we should develop the energy-saving, environmental friendly scheduling model which is based on grid security and reliability, and provides strong support for resource saving and environmental amiability society.

The power-dispatching agency of the USA, Australia, and other developed countries mainly do the dispatch based on the results of power market clearing and long-term contract trade of auxiliary service, the power consumption and pollutant emission reflected in the market electric price. So far, there is no application of power dispatching reporting real-time monitoring with power consumption, pollutant emission and heating.

The thermodynamic system employs the principle of thermal power conversion, and then forms a closed thermodynamic cycle curve by the characteristics of equipment quality, the structure and parameters of thermodynamic system [2–3]. Zhonghe Han [4] deduces the thermodynamic cycle of water vapor, and receives the impact of parameters on standard coal consumption rate of power generation. Napalko [5] employs theories of partial differential equations to calculate the efficiency of extraction steam, and makes the concept of thermal system of power plant much clearer.

In this chapter, a flow dimension is added to the traditional thermodynamic cycle diagram with two-dimensional enthalpy entropy, and a three-dimensional space model of enthalpy entropy flow is constructed. The comprehensive value of energy saving and environmental protection is calculated by adjusting the related environmental factors. The classic evolutionary programming is designed to receive the optimal power unit output distribution scheduling, and the efficiency of the energy-saving scheduling is verified by related examples.

This chapter is organized as follows. Section 37.2 is the researches on state space model of thermodynamic system. Section 37.3 describes the application of dispatching of energy saving and environmental protection. Section 37.4 discusses the experimental results of algorithm, followed by the conclusion in Sect. 37.5.

37.2 State Space Model of Thermodynamic System

37.2.1 Basic Concept

During a long optimization time, the units which contain the related equipment have been behaved well functionally. So the set of all possible stable operation units are called state space of units. Each factor is the satiation of stable operation, which contains various environmental temperatures, coal type for combust ion and the entire possible electric load.

The state space of thermodynamic system is a subset of the power plant space, and only contains main equipments which the circular material flow past, such as boiler heating surface, steam valve, condenser and feed water heater. Therefore, the corresponding factors for the elements of thermodynamic system state space are as follows: cycle efficiency, average endothermic temperature, and relative internal efficiency.

37.2.2 Physical Meaning of Model

A flow dimension is added to traditional enthalpy entropy thermodynamic cycle diagram, and the three-dimensional space model with enthalpy entropy flow is constituted. For example, at the current station, the coordinate points which represent the mainstream upward node state of circular material thermodynamic system are linked together, and it forms a closed curve.

The physical meanings of thermodynamic system state space model are as follows: the closed curves at three-dimensional spaces represent the experience of thermodynamic cycle. The trend of the blue curve approaches to the red by adjusting various control factors, then makes the actual state close to the optimal state, and finally realizes the optimization operation.

37.2.3 Model Description

Suppose the X^j is the state space which consists of N nodes in thermodynamic system, then X^j is the j-th factor of thermodynamic state space, and it represents j stable operation state.

$$\bar{X} = X^j (j = 1, 2 \cdots \infty) \tag{37.1}$$

Let $x_i^j (i = 1, 2, \cdots n)$ be the thermodynamic state node i at the j station. By the three-dimensional space descriptions, x_i^j is a vector, and the three components respectively mean the enthalpy, entropy and flow. The expression of thermodynamic system is as follows:

$$X^j = \left\{ X_i^j \right\} (i = 1, 2, \cdots n). \tag{37.2}$$

The derivatives of x_i^j are actual control elements that affect the state of node i ,then the boundary conditions and the related character can be described quantitatively. Let $\left\{ dS_k^j \right\} (k = 1, 2, \cdots m)$ mean control factor set, $ValT$ be the economy of current state, and $BestT$ perhaps be the economy of best state. So there are the following expressions:

$$ValT = f(x) = \frac{q_{cp}}{29.31 \cdot \gamma_b \cdot \gamma_p \cdot (1 - \xi_{ap})}, \tag{37.3}$$

where q_{cp} is the heat consumption of turbine, γ_b is the efficiency of the boiler, γ_p is the pipeline efficiency, and ξ_{ap} is the power consumption rate, form (37.3) is to calculate the actual coal consumption that means the current state economy.

$$dValT = ValT - BestT \tag{37.4}$$

Form (37.4) is adopted to calculate the difference between the best state and the current state.

$$dValT = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot dx_i \tag{37.5}$$

Form (37.5) describes the values of varying events that generate coal consumption changing.

$$dx_i = \sum_{j=1}^{m} \frac{\partial g_i}{\partial y_i^j} \cdot dy_i^j \, (i = 1, 2 \cdots n) \tag{37.6}$$

Form (37.6) is employed to describe the matters for the coal consumption varying events.

Based on the above analysis, we can obtain the calculation of generating cost and dynamic history optimal conditions searching. Via establishing the ration relationship between control goals and guiding factors, it can lay a foundation for dispatching of energy saving and environmental protection.

37.3 Application of Dispatching of Energy Saving and Environmental Protection

37.3.1 Objective Function

In order to achieve the requirements of grid security and the dispatching of energy saving and environmental protection, the objective functions were established with the operation data, and they are used to quantitatively describe the factors of scheduling priority, including coal consumption of power generation, operation state of desulphurization system, the concentration of dust and nitrogen oxides, the deviation of return water temperature and required value, adjustment margin of unit output and the load factor. We combined above and obtained the following four factors: coal consumption, energy-saving potential, quantity of pollutant emissions, return water temperature, and then received unit comprehensive value of energy saving and environmental protection by different weight values.

Objective functions:

$$\begin{aligned} \min \; F(\mu_i^t, P_i^t) = \sum_{t=1}^{T} \sum_{i=1}^{G} (&\alpha^e \cdot \mu_i^t f_i^e(P_i^t) + \alpha^s \cdot \mu_i^t f_i^s(P_i^t) \\ &+ \alpha^{ep} \cdot \mu_i^t f_i^{ep}(P_i^t) + \alpha^h \cdot \mu_i^t f_i^h(P_i^t)) \end{aligned} \tag{37.7}$$

$$Constra\,\text{int}\,conditions:\begin{cases}\text{power balance} P=\sum_{i=1}^{G} P_{\text{i}} \\ \text{output condition} P_{\text{imin}} \le P_i \le P_{i\max} \\ \text{slope condition}\left|P_i^{\text{t}}-P_i^{t-1}\right| \le dP_i\end{cases} \tag{37.8}$$

Form (37.7) $F(\mu_i^t, P_i^t)$ is the exponent of energy saving and environmental protection dispatching; T is the sum of time frame; G is total number of generators; P_i^t is the load power of i-th generator at the time t; $f_i^e(P_i^t)$, $f_i^s(P_i^t)$, $f_i^{ep}(P_i^t)$, $f_i^h(P_i^t)$ are the characteristic function of above four factors; $\alpha^e, \alpha^s, \alpha^{ep}, \alpha^h$ are the weight values and the sum is one; μ_i^t is the station of generator, $\mu_i^t = 1$ is running and $\mu_i^t = 0$ means stopping.

37.3.2 Classic Evolutionary Programming

The classic evolutionary programming is proposed by L. J. Fogel [6–8] to solve various optimization problems. As for the application of energy saving and environmental protection dispatching [9], firstly, the output unit based on load prediction curve, current load of each unit and interval value of ACE is initialized, and then the comprehensive value is calculated; secondly, the population is generated and the current best fitness individual is held until the time equals T, and then the optimal distribution scheme is obtained; finally, the optimal apportioning cost of each unit is output.

37.3.3 Flow of Scheduling Data

In this chapter, the dispatching data processing consisted of data acquisition, primary treatment, data transmission, data storage, data after treatment and exhibition. The front-end data acquisition servers (substation) which are set up at power stations are used to load the real-time data of unit operation information system (SIS), and the data transmitted in dispatching communication network adopts 104 protocol to D5000 front-end system. At the same time, the substation server data are uploaded to the workstation which deploy at power station. The processed data can be shown by web browsing for grid side user, and network wiring diagram of Hefei power is depicted in Fig. 37.1. The operation data of power plant side upload to the provincial dispatching center, and write into the D5000 database by platform front-end processor.

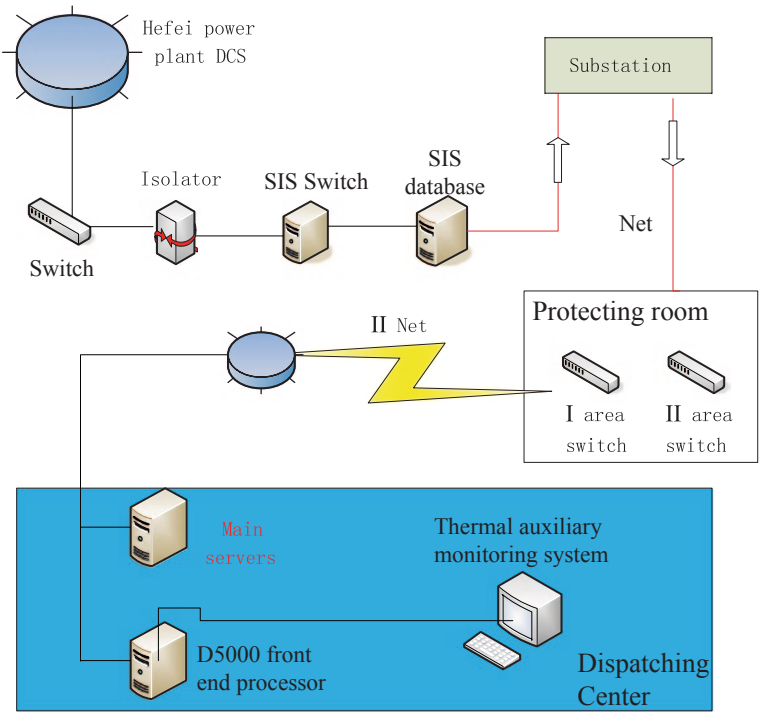


Fig. 37.1 Network of Hefei power

37.4 Analysis of Examples

The applications of grid dispatching of energy saving and environmental protection are based on thermodynamic system model with flow dimension; we show the examples with fifth unit of Hefei power plant to effectively analyze by the following three steps: original data analysis, accuracy analysis, and efficiency analysis.

37.4.1 Original Data Analysis

The fifth unit of Hefei power plant is a 660-MW water-cooled generator. We choose the data from 20 April 2013to 05 June 2013, and the total number of data is 12,161 points. Depending on the history data of normal stable experiments, the data is chosen which satisfied inspection standard, given in Table. 37.1.

Table 37.1 Constitute of original data

Sample unit	Inspection standard		Total number	Effective data	Rate
	Load	Coal consumption			
Hefei power plant #5 unit	(50,1000)	(280,400)	12,161	12,127	99.72%

Table 37.2 Correctness analysis of unit data

Testing data	Test criterion	Over lower limit ratio	Standard range ratio	Over upper limit ratio	Average value
Coal consumption	(290,320)	2.21	80.04	17.75	303.54
Boiler efficiency	(0.88,0.93)	0.1	99.9	0	92.45
Cycle efficiency	(0.49,0.55)	2.94	97.05	0.01	53.83
Relative internal efficiency	(0.85,0.90)	2.19	60.62	26.18	89.73
Auxiliary power ratio	(0.05,0.11)	0.01	99.97	0.02	5.10
High-pressure cylinder efficiency	(0.72,0.87)	18.11	81.89	0	79.10
Middle-pressure cylinder efficiency	(0.88,0.96)	0	82.67	17.22	94.16
Low-pressure cylinder efficiency	(0.82,0.95)	0	99.98	0.02	91.89

37.4.2 Accuracy Analysis

The effective data are selected by the conditions of energy consumption, then the over lower limit ratio, standard range ratio, over upper limit ratio, and the average value are calculated as shown in Table 37.2.

37.4.3 Efficiency Analysis

The unit efficiency of energy saving and environmental protection dispatching mainly contains four aspects: coal consumption, boiler efficiency, turbine heat consumption, and auxiliary power ratio. The details of analysis are as follows.

The coal consumption for fifth unit is shown in Fig. 37.2; we can see that the average value is 308.1807 g/kWh, and exceed 4.2442 g/kWh compared to the history optimal. Figure 37.3 depicts the boiler efficiency of the unit, in which the average statistics value is 92.0803 %, approaching to the history optimal 92.2443 %. From Fig. 37.4, the value of turbine heat consumption is 7743.4963 and 54.2989 kJ/kwh more than history. Figure 37.5 describes the auxiliary power ratio, and the average is 5.4644 %.

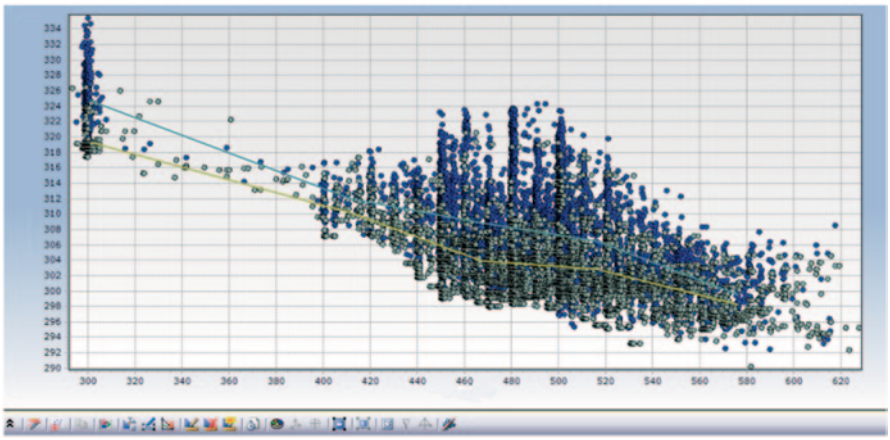


Fig. 37.2 Coal consumption

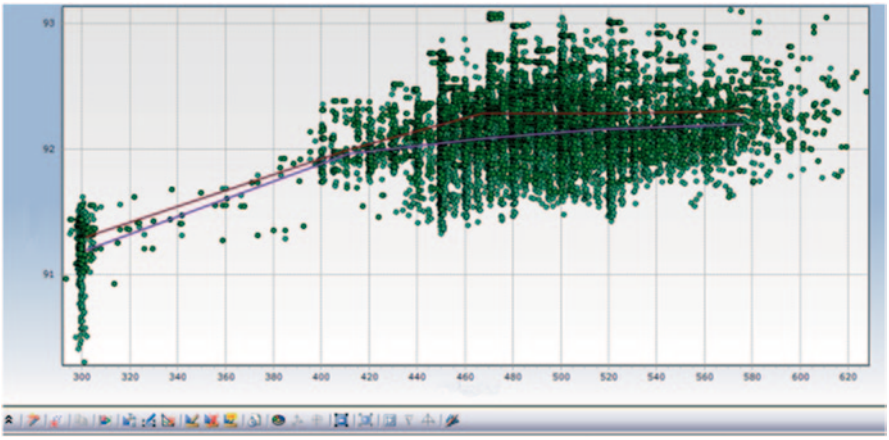


Fig. 37.3 Boiler efficiency

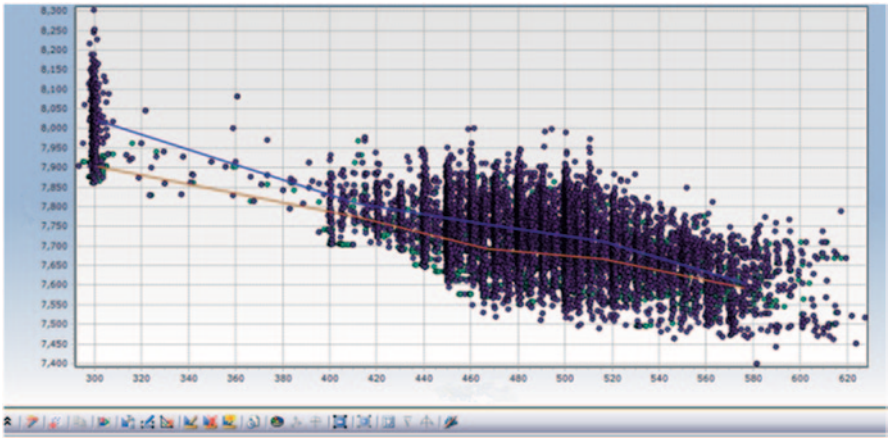


Fig. 37.4 Turbine heat consumption

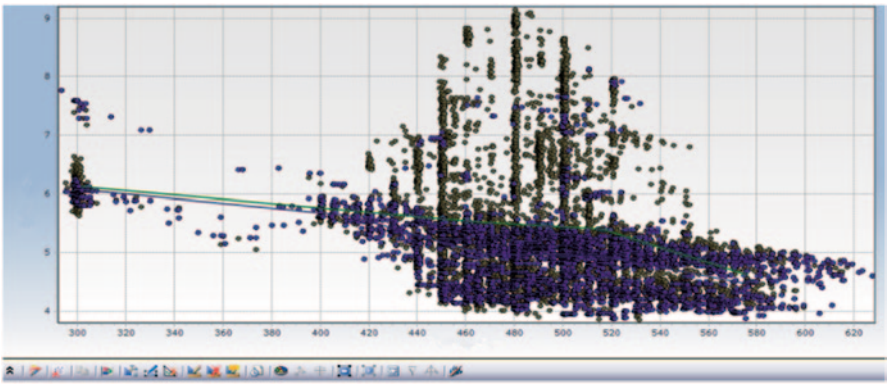


Fig. 37.5 Auxiliary power ratio

37.5 Conclusion

A novel thermodynamic system model with flow dimension is proposed to solve the problem of grid energy saving and environmental protection dispatching, and then the optimal output distribution scheme is obtained by comprehensive index value and the generation of classic evolutionary program algorithm. The example analysis shows the possibility of developing the energy-saving, environmental friendly scheduling model based on grid security and reliability, providing a strong support to construct resource saving and environmental amiability society.

References

1. Liu Z. China power and energy. Beijing: China Power Press; 2012. p. 60–88. (In Chinese).
2. Vahidi, Tavakoli, etc. Determining parameters of Turbine's model using heat balance data of steam power unit for educational purposes. Power Syst. 2007;22(4):1547–53.
3. Mandal KK, Basu M, Chakraborty Nï¼¼Particle swarm optimization technique based short-term hydrothermal scheduling. Appl Soft Comput. 2008;8(4):1392–9.
4. Han Z, Yang K. A Thermodynamics method of online calculation steam turbine exhausts enthalpy. Power Eng. 2004;24(3):356–9. (In Chinese).
5. Napalkov. On the theory of the linear Partial differential equations with variable coefficients. Dokl Akademii Nauk. 2004;397(6):748–50.
6. Sahoo PK. Exergoeconomic analysis and optimization of a cogeneration system using evolutionary programming. Appl Therm Eng. 2008;28(9):1580–8.
7. Fogel DB. Evolving artificial intelligence. San Diego: University of California; 1992. p. 48–96.
8. Yao X, Liu Y. Evolutionary programming made faster [J]. IEEE Trans Evol Comput. 1999; 3(2):82–102.
9. Liang X, Yue L-h, Xiong Y, Cheng W-j. On the analysis of evolutionary programming with self-adaptive cauchy operation. Chin J Electr. 2012;21(2):309–12.

Chapter 38
Targeted Comprehensive Strategy for Lightning Protection and Its Application in the Anhui Power Grid

Xuefang Tong, Jiahong Chen and Shiling Zheng

Abstract In order to instruct more scientific and effective lightning protection methods, a targeted comprehensive strategy for lightning protection from surface to line to point by synthesizing operating experience, lightning zones map, lightning risk assessment, and treatment technology is proposed based on the practical management demand and differentiation idea of lightning protection. The strategy is applied to the Anhui power grid. Firstly, the characteristics of historical lightning faults are summarized; secondly, the lightning zones maps in the overall province and along transmission lines are plotted to determine regions with strong lightning activities; thirdly, according to the lightning risk assessment and the treatment technology, the risk grades as well as its determination for typical transmission lines are evaluated, and several reform schemes are made with their technology and economy as definitely given. The related results have been applied to the lightning protection work in the Anhui power grid.

Keywords Lightning faults · Differentiation lightning protection · Lightning zones map · Risk assessment · Reform scheme

38.1 Introduction

The transmission lines spread out widely in the open, and the lightning is one of the most severe threats to the safe and stable operation of these lines. According to the latest statistics of State Grid Corporation of China (SGCC), lightning is the main reason of lines' trip-out, accounting for about 40 % [1]. With the further

X. Tong (✉)
China Electric Power Research Institute, 430074 Wuhan, China
e-mail: txfang82@163.com

J. Chen
State Grid Electric Power Research Institute, 430074 Wuhan, China

S. Zheng
Anhui Electric Power Research Institute, 230022 Hefei, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_38

development of the social economy, the power grid capacity will expand, lines and towers will be denser and the geographical situations along lines will be more serious, which may bring more lightning faults and demand more effective lightning protection method.

Usually, the lightning protection primarily adopts some technical measures for key sections; but it is hard to find out the effective protection sections by only relying on experience because of strong randomness of lightning. If the protection target is not clear, the technology and economy of adopted measures will be poor.

With the accumulation of lightning monitoring data, statistical method of actual lightning parameters in specific region is proposed [2–3]. Then, statistical lightning parameters are gradually adopted in lightning performance analysis [4–6]. Differentiation risk assessment method that evaluates the lightning flashover risk for every span or every tower was proposed in 2009, which took all kinds of influent factors into account, such as lightning activity, topography, line structure, insulation configuration, and protective measure [7]. The differentiation risk assessment method is effective, which can provide the lightning performances of different sections and figure out the key sections with corresponding reasons.

It is unpractical to adopt differentiation risk assessment elaborately for every line. This paper proposes a targeted comprehensive lightning protection strategy from surface to line to point based on management need and differentiation protection idea. According to the strategy which has been applied in the Anhui power grid, the operating experience of lightning protection is summarized, the distribution characteristics of lightning zones are researched, the lightning performance and the risk grade of a typical line are evaluated, and the available reform schemes are drawn up for an example.

38.2 Targeted Comprehensive Strategy for Lightning Protection

In different regions, the lightning activities characteristics are different, as well as the lightning performances of lines, the operation and management levels, the protective measures, etc; in this sense, the idea of “differentiation protection” and “comprehensive protection” should be used to guide scientific lightning protection work.

Under the guidance, it is recommended that in combination with the operating experience, the decision of overall layout of lightning protection should be based on the lightning zones map of the overall area, the decision of key protection sections for ordinary lines should be based on the lightning zones map along the lines, and the decision of key protection points and reform schemes should be based on the lightning risk assessment and treatment analysis for important lines or sections in order to realize the targeted comprehensive strategy of “surface–line–point,” whose process is shown in Fig. 38.1.

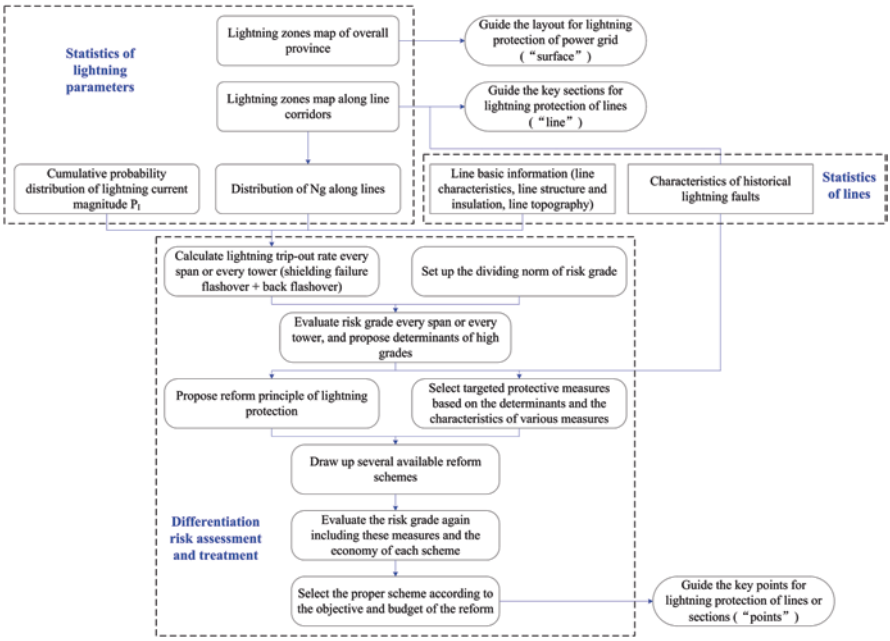


Fig. 38.1 Process of the targeted comprehensive lightning protection strategy from surface to line to point

1. Characteristics of historical lightning faults
   The operating experience is a very important reference to verify various measures and basic theories. The characteristics of historical lightning faults and the operating effects of protective measures should be summarized timely, orderly and gradually. It is better to focus on the characteristics that are likely to cause lightning accident in the lightning protection work.
2. Technology of lightning zones map
   The lightning is the source of lightning accident, and the lightning zones map is the important basis to guide the lightning protection layout in the targeted design and reform. There are two kinds of lightning zones maps: the lightning zones map of overall province and the lightning zones map along line corridors. In the planning of power grid, it should avoid regions where the lightning activities are intense in the lightning zones map of overall province as far as possible. In the design of new lines and the reform of operated lines, the protection should be focused on or inclined to regions with strong lightning activities in lightning zones maps along line corridors.
3. Technology of differentiation risk assessment and treatment
   As to important transmission lines or sections, more elaborate technology of lightning risk assessment and treatment is further recommended to find key points, considering various influent factors and the ability of resisting lightning. Key points with high lightning risk and corresponding weak reasons are gained,

based on which several feasible reform schemes are made as well as their technology–economy rates are estimated.

38.3 Characteristics of Historical Lightning Faults

The lightning trip-out faults data of the power grid in Anhui from 2003 to 2008 are collected, which are the basis of mastering the characteristics of historical lightning faults. There are 101 lightning trip-out faults available on 220 and 500 kV transmission lines. The trip-out rate is 0.18 flashovers/100 km/year for 220 kV lines and 0.20 flashovers/100 km/year for 500 kV lines. The former exceeds the norm of SGCC especially in 2003 and 2007, while the latter is within the norm [8]. Based on these data, some characteristics of lightning faults are found as follows:

In a year, the lightning faults mainly occur in summer, accounting for 84.1%. In 1 day, the lightning faults mainly occur from afternoon to the first night, accounting for 55.1% between 14:00 and 21:00.

In space, the lightning faults are concentrated relatively on Wuhu and Tongling.

Shielding failure flashover is the main fault type, accounting for 74.3% for 220 kV lines and 87.1% for 500 kV lines.

In terrain, along the slope and top of mountain where more lightning faults occur are the key terrains which deserve close attention.

The phenomenon of double-circuit simultaneous tripping is not obvious, accounting for 1.4% only.

Straight line tower is the main fault tower style, accounting for 74.4%. The tower styles of ZM in one-circuit lines and PSL in double-circuit lines are main styles with more lightning faults in 220 kV lines.

Towers with big shielding angle should be focused on, as well as those with a height greater than 60 m. There are 75.0% fault towers higher than 60 m even with negative shielding angle.

38.4 Lightning Zones Maps

38.4.1 Lightning Zones Map of Overall Province

Based on the grid method [2], the statistical region is divided into a number of grids with 0.01°×0.01°, and then the ground flash density Ng is analyzed in each grid. The lightning zones map is based on Ng, taking 4, 5, 6, 8 flashes/km^2/year as the dividing points to divide the whole area into five grades labeled as I~V, respectively from low to high, painting different colors in different grades. The lightning zones map of Anhui is shown in Fig.38.2. It is found that lightning activities in most regions are strong with most of them above grade III and the average Ng of 5.6 flashes/km^2/year. The lightning activities are mainly concentrated in the south of Anhui, such as

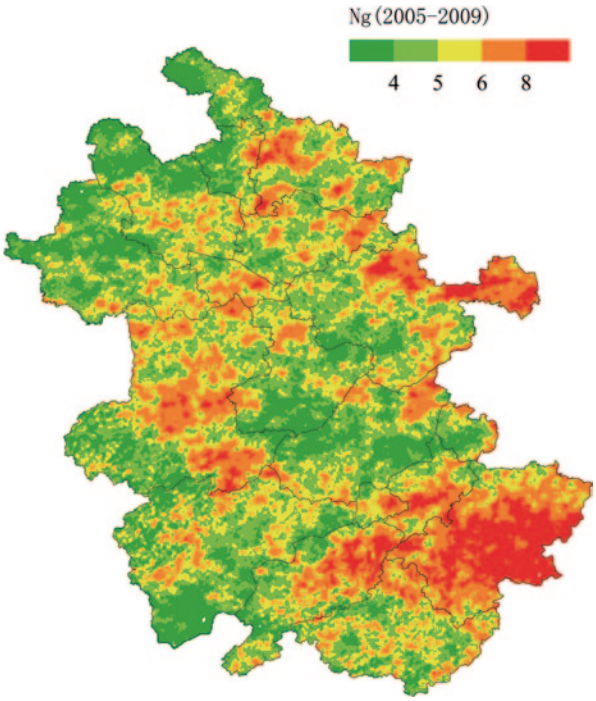


Fig. 38.2 Lightning zones map of Anhui Province

Xuancheng, Wuhu, Tongling, southeast of Anqing, east of Chizhou, and northeast of Quzhou. The lightning activities in the northwest are relatively poor.

The lightning zones map of overall province has been important data in lightning protection work in the Anhui power grid. Before the construction of new facilities such as substation and transmission line, the lightning activity intensity in the proposed location will be analyzed based on the map to avoid the thunderstorm area as far as possible.

38.4.2 Lightning Zones Maps along Line Corridors

It is more exact to consider lightning parameters along the line corridors for operated lines. Similar to the grid method, divide a line corridor into several grids, and then analyze Ng in each grid to form the lightning zones maps along the line corridors, still taking 4, 5, 6, 8 flashes/km^2/year as the dividing points, as shown in Fig. 38.3 for an example.

Analyze statistically 265 lines of 220 kV and 59 lines of 500 kV in the Anhui power grid, with the corridor width 4 km. Lightning activities along all line corridors are mainly in grade II-IV, and the ratio of each grade from low to high levels is 11.8, 26.2, 29.7, 24.8 and 7.5%, respectively. The ratio of each grade of 220 kV lines is similar to that of 500 kV lines.

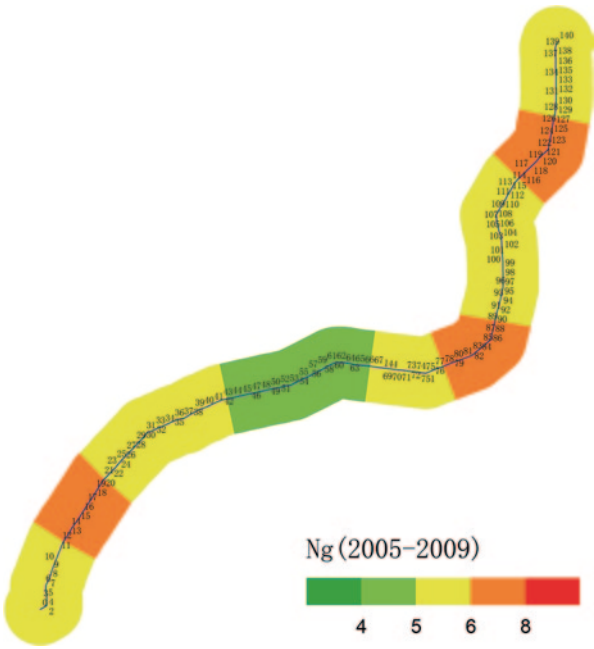


Fig. 38.3 Lightning zones map of FC line

Account the numbers of lightning faults and towers in each grade, and calculate the fault ratio (the ratio of lightning faults to towers) in each grade. The ratio of each grade from low to high levels is respectively 0.13, 0.18, 0.20, 0.36 and 0.64%. Visibly, with the grade level increasing, the probability of lightning fault increases gradually. The lightning activity level has a significant effect on the risk of lightning strikes.

Based on the above lightning zones maps of 324 lines, the lightning zones book is produced in the Anhui power grid to instruct the lightning protection reform of ordinary lines for every supply company. Reforms are inclined to sections with high grade level.

38.5 Differentiation Risk Assessment and Treatment for Lightning Protection of a Typical Transmission Line

38.5.1 Differentiation Risk Assessment for Lightning Protection

Based on the statistical results of lightning parameters such as Ng and cumulative probability distribution of the lightning current magnitude P_I, the Monte Carlo method and electric geometry model are adopted to evaluate the lightning flashover risk of typical transmission line. Taking 500 kV YC line as an example, the lightning trip-out rate results along the line are shown in Fig. 38.4. The shielding failure flashover

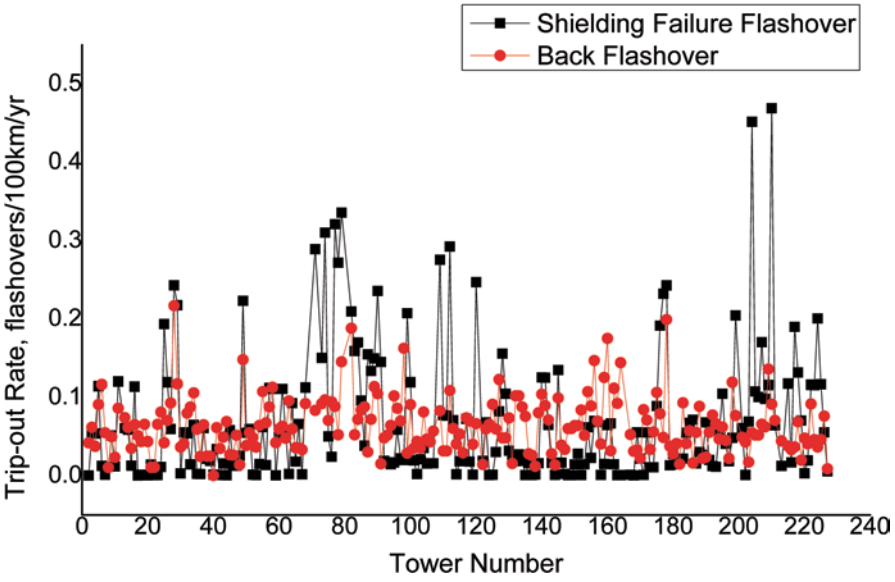


Fig. 38.4 Calculated results of lightning trip-out rate along the YC line

rate of each section is mainly in 0.01–0.30 flashovers/100 km/year, and the back flashover rate of each section is mainly in 0.02–0.16 flashovers/100 km/year.

Convert the operating norm of lightning trip-out provided by SGCC into Sr through actual lightning activity level, and set Sr as the trip-out rate benchmark for risk assessment [8]. Take 0.5Sr, 1.0Sr, and 1.5Sr as the dividing points to divide the risk level into four risk grades labeled as A, B, C and D respectively from low to high. The distribution of risk grade of YC line is shown in Fig. 38.5, where 10.8% are above risk grade C in shielding failure flashover, and 29.1% are above risk grade C in back flashover. The main reasons for the less overall flashover risk include that shielding angles of towers are negative, spans between towers are small (<500 m), the line is located in the plain where ground tilt angle is small (<2°), and

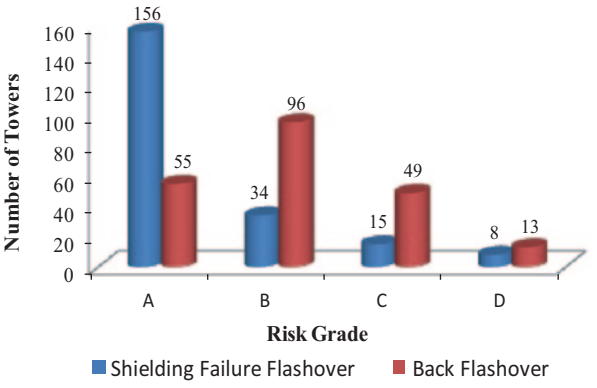


Fig. 38.5 Distribution of risk grade along the YC line

the line is new which is put into operation in 2007; therefore, it should be focused on towers with higher risk in the next lightning protection reform.

One shielding failure flashover fault happened on tower 127# in 2007 actually. The actual lightning trip-out rate is 0.186 flashovers/100 km/year, which is similar to the theoretical calculation result, both in risk grade B. In the three years of operation, the lightning flashover risk level in the fault year (in risk grade C) is higher than the other 2 years (in risk grade B and A separately). The risk assessment results are highly in conformity with the operating experience.

38.5.2 Differentiation Treatment for Lightning Protection

The purpose of lightning protection is to reduce the lightning trip-out rate. Operation experience shows that the shielding failure flashover was the main fault type for 500 kV lines as analyzed before, and the only one lightning fault for YC line was caused by shielding failure; therefore, the reform schemes aim at reducing the shielding failure risk.

Based on the characteristics of historical lightning fault, the risk assessment results and influenced factors, reform towers and order are determined as below: (1) towers with lightning fault; (2) towers with lightning risk grade D; and (3) towers with lightning risk grade C.

Based on the operation experience and actual situation of the line, the controllable lightning rod and the line arrester are the main measures adopted in the reform. Controllable lightning rod is cheaper and effective for shielding failure. The effect of line arrester is better but more expensive, which is planned for towers with lightning fault, higher soil resistivity and higher risk grade. Two controllable lightning rods will be installed on both sides in each corresponding tower. The installation number and phase of line arresters depend on tower terrain and shielding failure probability of each phase. The overall soil resistivity is small along the line; thus, two line arresters will be installed for towers located on top of mountain, and one line arrester will be installed for towers located along the slope.

Four available reform schemes are presented based on the above principles, whose technology and economy are estimated as shown in Table 38.1.

Table 38.1 Technology and economy of various reform schemes for YC line

No.	Number of controllable lightning rods	Number of line arresters	Reach norm or not	Percentage of superior operating experience (%)	Economy/ Wan Yuan ¥
1	0	14	Yes	18.4	109.2
2	42	5	Yes	27.9	106.2
3	36	11	Yes	37.3	143.4
4	30	14	Yes	44.1	151.2

38.6 Conclusion

1. Based on the demand of operating management and the idea of differentiation lightning protection, the targeted comprehensive strategy of "surface–line–point" for lightning protection is proposed, which synthesizes operating experience, lightning zones maps, differentiation risk assessment and treatment technology.
2. The targeted comprehensive strategy for lightning protection is applied to the Anhui power grid. Firstly, the characteristics of historical lightning faults and protection level of transmission lines are summarized with some typical influent factors proposed. Secondly, the lightning zones maps in the overall province and along all 324 transmission lines are plotted based on Ng to determine sections with strong lightning activities. Thirdly, the lightning performance of typical transmission line YC is evaluated, pointing out high flashover risk points. According to the assessment results, in combination with the operating experience, four available reform schemes are established as well as their technology and economy to provide strong support for the decision-making in terms of final implementation plan for governor.
3. The lightning zones maps and lightning risk assessment results can effectively reflect the difference in lightning activity level and lightning protection performance, in good agreement with the operating experience.
4. The process and objective of the targeted comprehensive strategy are clear, and it has been applied to the lightning protection work in Anhui.

References

1. State grid electrical power research institute. Survey and assessment of lightning protection level and strategy of power transmission equipments in state grid corporation. Technical Report, Wuhan; 2010. p. 3–8. (In Chinese)
2. Chen J, Feng W, Wang H, et al. Statistical method of lightning parameters. High Vol Eng. 2007;33(10):6–10. (In Chinese)
3. Orville RE, Huffines GR. Cloud-to-ground lightning in the United States: NLDN results in the first decade, 1989~1998. Mon Weather Rev. 2001;129(5):1179–93.
4. Gu Sh, Chen J, Chen W, et al. Evaluation method of the time-space-difference of lightning protection performance of transmission lines. High Volt Eng. 2009;35(2):294–8. (In Chinese).
5. Shim EB, Woo JW, Han SO, et al. Lightning characteristics in Korea and lightning performance of power systems. Transmission and Distribution Conference and Exhibition 2002: Asia Pacific, IEEE/PES; 2002. 1:534–9.
6. Xie Y, Xie W, Gu Sh, et al. Lightning performance analysis of shielding failure for 220 kV transmission lines with double circuit on the same tower using lightning parameter statistics. High Volt Eng. 2009;35(11):2657–62. (In Chinese).
7. Chen J, Lv J, Qian Zh, et al. Differentiation technology and strategy of lightning protection for transmission lines. High Volt Eng. 2009;35(12):2891–902. (In Chinese).
8. State Grid Corporation of China. Lightning protection management regulations of power transmission equipments of state grid corporation of China. Technical Documents, Beijing; 2005. p. 8–8. (In Chinese).

Chapter 39
Saturated Load Forecasting Based on Nonlinear System Dynamics

Haihong Bian and Xindi Wang

Abstract Saturated load of a power system is the key index for the local grid planning, which identifies the final scale of a power system. Due to the long time span and sensitivity to economic factors, the precision and reliability of the direct saturated load forecasting (SLF) are not satisfied. Therefore, this chapter mainly proposes a novel SLF model derived from the saturated economy forecasting (SEF), based on nonlinear system dynamics. A practical case was investigated according to the real economic and load data of Fujian province, China. The method proposed was proved reliable, with a consistent result but more flexibility and extension to the per capita electricity consumption (PCEC) method.

Keywords Power system · Nonlinear system dynamics · Saturated load forecasting · Saturated economic forecasting · Per capita electricity consumption

39.1 Introduction

Saturated load is the load limit when the increasing loads obviously grow slower or even stop growing after experiencing the rapid growing phase, steady growing phase and the transition phase. Its value is limited by the local environment, resources, and grid planning. It is also identified as one of the key indicators of the regional grid planning, which decides the final size of the grid.

Unlike the traditional long-term load forecasting, the saturated load forecasting (SLF) is more complex not only because its time span is far longer but also because the condition of economy and resources and even saturated economic forecasting (SEF) must be considered together as a whole. Domestic and foreign researchers

H. Bian (✉) · X. Wang
School of Electronic Power Engineering, Nanjing Institute of Technology,
211167 Nanjing, China
e-mail: bianhh.njit@gmail.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_39

usually study it with historical data by using traditional methods such as time series, regression analysis and ANN model [1–5].

If saturated load is predicted after the saturated economic, the result may be more convincing. Hence, the system dynamics, a method that is widely applied in economic forecast and power system load forecast, is used in this chapter to predict the saturated load.

39.2 System Dynamics

System dynamics (SD), first proposed in the fifties of the last century by Massachusetts Institute of Technology Professor Forrest, is a branch subject to analyze the feedback information of a system [6, 7]. It is a practice-oriented modeling approach, which can play the advantages of the complex computing power and combine the effective of scientific theory, empirical knowledge, and expert theory together to solve complex nonlinear systems through flow diagrams [8].

In this chapter, we use SD to forecast saturated load based on SLF with the economic data and the historical load data of Fujian province of China.

39.3 Saturated Economy Forecasting Model

39.3.1 Economic Forecasting Method

Using traditional methods separately often leads to a large error, though the prediction accuracy may be coincidently satisfactory in very few cases. So a combination forecasting method is used. The model combines different forecasting methods together appropriately so as to improve the forecasting accuracy. We apply the specific optimal linear combination of prediction methods on the model, which is based on the linear regression equation:

$$Y = a + \sum b_i X_i \tag{39.1}$$

According to the principle of least squares method, we can make the error between the forecasting and actual values smallest. There are three forecasting methods used in the model: time series ARMA model [9], curve estimation model with SPSS software [10], and grey prediction model GM (1, 1) [11].

39.3.2 Solow Economic Growth Model

The Solow neoclassical growth model is the cornerstone of modern growth theory. According to the form of Cobb-Douglas production function in the Solow model, we transform the basic equation into a logarithmic one. What is more, we expand the basic model in order to investigate the impact of factors on total factor productivity (TFP) such as urbanization, opening up, and technological progress:

$$\begin{aligned} \ln Y(t) &= C + a_1 \ln P(t) + a_2 \ln H(t) + a_3 \ln CR(t) \\ &+ a_4 \ln TR(t) + a_5 \ln FDI(t) + a_6 \ln cons(t) \\ &+ a_7 \ln cons^2(t) + \varepsilon(t), \end{aligned} \tag{39.2}$$

where Y represents the gross domestic product (GDP) and K and H represent physical capital stock and human capital stock, which are calculated based on 1978 constant prices. CR represents urbanization rate; FDI is actually utilized foreign direct investment; TR represents dependence of foreign trade, which means total imports and exports share of GDP; $cons$ is the final consumption rate; ε is the random error term. For final consumption $cons$, it is assumed that it will nonlinearly affect the economic efficiency if its percentage of GDP is too high and too low. Accordingly $cons$ and its square term are included in the model so as to test its hypothesis.

Finally, we apply the least square method to estimate the above model. The elements' flexibilities are shown in Table 39.1. All the models and parameters in this chapter have been tested.

For the forecasting of industrial structure, we use the compositional data model. Three major industries GDP data of Fujian province from 1978–2008 are recognized as the compositional data in this model so that industrial structure can be established. According to the method, the compositional data is made a simple nonlinear transformation [12, 13]. And then, after the proportion of second industry corresponds to θ_2^t and the proportion of third industry corresponds to θ_3^t, the time series of θ_2^t and θ_3^t will be fit with the Fourier function curve.

Table 39.1 The flexibility of elements in economic growth in Fujian province

Variable	Flexibility	Standard error	t statistic	Pr (>|t|)
P	0.35926	0.05628	6.383447	2.01E-06
H	0.64074	–	–	–
TR	0.120161	0.042765	2.809832	1.02E-02
FDI	0.058805	0.012708	4.627486	0.00013
CR	0.518955	0.227728	2.278838	3.27E-02
Cons	−1.27627	0.253962	−5.02543	4.95E-05

Table 39.2 The flexibility of elements in economic growth in Fujian province. The forecast results of industrial structure during 2010–2050 in Fujian province

Year	Growth rate	First industry	Second industry	Third industry
2010	12.48			
2015	9.93	7.8	52.1	4.01
2020	7.57	6.5	52.0	41.5
2025	5.97	5.7	50.6	43.7
2030	4.81	5.0	48.2	46.8
2035	3.92	4.5	45.1	50.4
2040	3.19	4.0	42.2	53.8
2045	2.58	3.7	39.9	56.4
2050	2.05	3.7	38.8	57.5

With the help of the fitted equation, θ_2^t and θ_3^t are predicted based on historical data. Meanwhile, the industrial structure can also be acquired after the antitransform of θ_2^t and θ_3^t. The results are shown in Table 39.2.

As we have the predictive value of all the motivation, the economic growth model can be used to forecast the economic growth in Fujian province, as shown in Table 39.2.

39.4 Saturated Load Forecasting Model

39.4.1 System Dynamics Model

The most important for predicting saturated load is to select main variables which have great influence on saturated load for the model. Traditional power load forecasting methods are influenced by a number of factors such as GDP, industrial structure, urbanization rate, and energy intensity [14–16]. It can increase the reliability and predictive ability of the model if we consider more factors. Some researchers advocated dividing the research into various parts according to the differences between industrial and residential electricity. In this way, the mutual restriction and interrelated features among these factors can be reflected by taking into account the complexity and universal applicability of the load forecasting model. In this chapter, the national policy and socioeconomic issues should be reflected in the factors. What is more, the model needs to be a certain degree of complexity and, on the other hand, is able to consider the experts' experience so that the model will be easy to use.

In this chapter, we use the Vensim PLE software to model and simulate. The SD model is shown in Fig. 39.1.

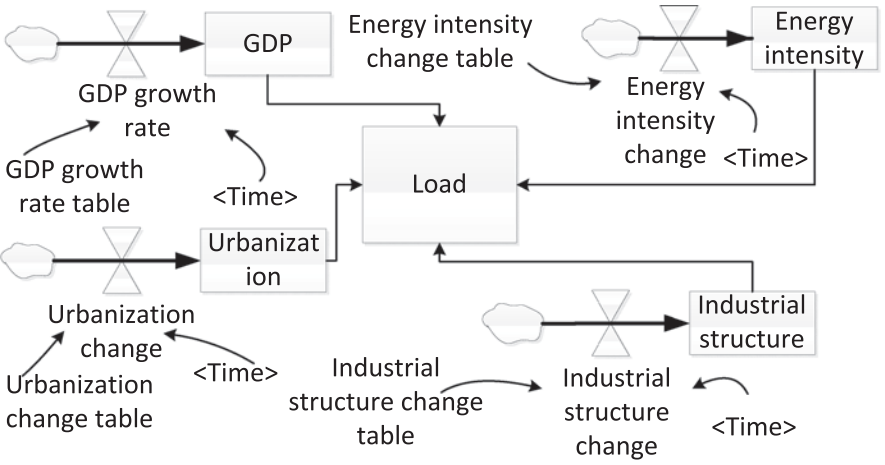


Fig. 39.1 The model of saturated load forecasting

The model equations:

1. Load solving equation:

$$L = (E_{GDP}W_{GDP} + E_{UR}W_{UR} + E_{IS}W_{IS} + E_{EI}W_{EI}) / (W_{GDP} + W_{UR} + W_{IS} + W_{EI}), \tag{39.3}$$

where L is the electric load, E_{GDP} is the GDP factors, W_{GDP} is the weight of the GDP factor, E_{UR} is the urbanization rate factor, W_{UR} is the weight of the urbanization rate factor, E_{IS} is industry structural factor, W_{IS} is the weights of industry structure factor, E_{EI} is the energy intensity factor, and the W_{EI} is the weight of energy intensity factor.

2. State equation (e.g., GDP factors):

$$GDP.K = GDP.L + DT * G_{GDP}, \tag{39.4}$$

where K is the value of the next point and L is the value of the current time point. G_{GDP} is the growth rate of GDP.

3. The auxiliary equation (e.g., GDP factor):

$$G_{GDP} = GDP * R_{GDP} \tag{39.5}$$

$$R_{GDP} = T_{GDP}(t), \tag{39.6}$$

where T_{GDP} is the time table of the growth rate of GDP.

39.4.2 Data Collection

The model needs two kinds of data—one type has influence on the factors' value and the other impacts weights. Factor data can be obtained from the SEF model, while the weights come from the analytic hierarchy process.

39.4.2.1 Factors Data

Factors data are economy data. We should research on the saturated economic situation so that we can find out the detailed status of the future economic developing condition. We have managed to obtain the conclusions of urbanization rate, GDP, and industrial structure in the SEF model.

39.4.2.2 The Weights of Factors

It is crucial to get the weights of the factors. We can certainly obtain the weights through regression method, and it is simple to operate. But the weights getting from the method above can only reflect the relationship between these factors and loads during every historical stage. It is debatable whether they can be used to strike a saturated load.

The analytic hierarchy process is excellent at evaluating the severity of factors comprehensively. It firstly creates the weights from the perspective of mathematic, by comparing the importance of every factor and then calculates synthetically. Yao et al. described the calculation steps [17].

First of all, the judgment matrix is formed according to the importance of the factors, as shown in Table 39.3.

Calculate the largest eigenvalue and we get the corresponding eigenvector (0.9481, 0.2587, 0.1568, 0.0975), so that we can calculate the value of the consistency index *CI:*

$$CI = \frac{\lambda_{\max} - n}{n-1} = 0.0255. \tag{39.7}$$

Table 39.3 The judgment matrix of the main factors

	GDP	Urbanization	Industrial structure	Energy intensity
GDP	1	5	6	7
Urbanization	1/5	1	2	3
Industrial structure	1/6	1/2	1	2
Energy intensity	1/7	1/3	1/2	1

Table 39.4 The forecasting results

Years	Predictive value (MW)	Average annual growth rate (%)
2020	35,222	5.50
2030	54,700	4.40
2040	68,789	2.83
2045	84,116	2.08

When *RI* is chosen to be 0.90, the *CR* indicators are as follows:

$$CR = \frac{CI}{RI} = 0.0283 < 1. \tag{39.8}$$

Formula (39.8) proves that the calculation results meet the consistency demand, and the actual weight (GDP, urbanization, industrial structure, energy intensity) should be 0.9481, 0.2587, 0.1568, and 0.0975.

39.4.3 Load Forecasting Results

The model was built in Vensim PLE software, and the beginning time was set as 2010, and 5 years as a stage, then the forecasting was continued until the load growth was less than 2%. The results are partly shown in Table 39.4.

According to the final prediction results, the saturation load of Fujian province will be about 85,000 MW.

The weights are chosen according to the forecast periods. And they also can be changed with the different stage characteristic.

39.5 Per Capital Electricity Consumption Method

In this chapter, we also apply the per capital electricity consumption (PCEC) method to SLF in order to test the result of the SD method. We need to predict the population, PCEC, and maximum load hours or find out the relationship between electricity consumption and maximum load.

The modified exponential curve fitting method is the description for a special class of growth phenomena that the parameter grows rapidly at first and then gradually slows down, and finally stops growing at the limitation *k*. The model is expressed as:

$$y_t = k + ab^t (k > 0, a < 0, 0 < b < 1). \tag{39.9}$$

With the history data of Fujian province during 1980–2009, we can acquire the values of the parameters of the exponential model: $a=-1942.3547$, $b=0.9699$, $k=444.4369$. So the correction curve model equation is as follows:

$$y_t = 4444.4369 - 1942.3547 * 0.9699^t. \tag{39.10}$$

The model shows that the population of the year 2045 will be 42.15 million.

PCEC is fit with the logistic curve. The curve equation is as follows:

$$y = k / (1 + ae^{-bx}). \tag{39.11}$$

We can put the PCEC data of Fujian province during 1980–2009 into the above equations. As a result, we can get the estimated values of the parameters of the exponential model: $a=216.4$, $b=-0.1071$, $k=11{,}900$, i.e., per capita saturated electricity consumption is 11,900 kWh/people. The fitting results are shown in Fig. 39.2. With the results of population and PCEC, we can get the annual electricity consumption as 499.99 billion kwh.

In the load forecasting of the power system, we usually predict the maximum load utilization hours to study the relationship between maximum load and the annual electricity consumption. However, the maximum load hours and load characteristics closely relate to national policies and are generally more difficult to forecast accurately. Wei et al. proposed a model based on the theory of probability and

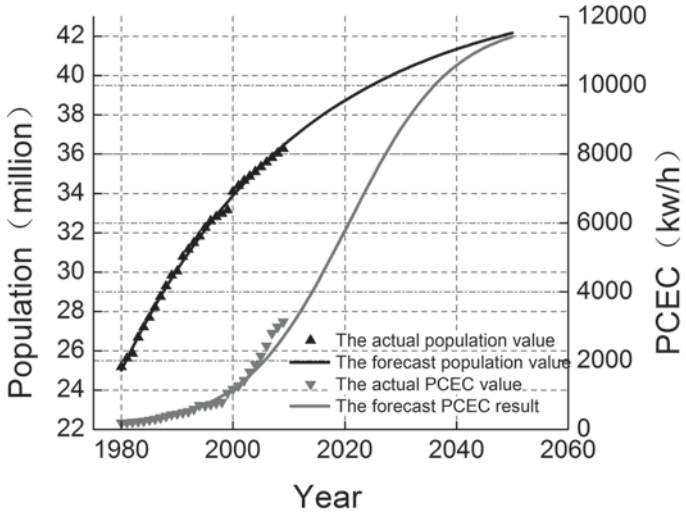


Fig. 39.2 The forecast results of population and per capita electricity consumption during 1980–2009 in Fujian province

statistics and establishes a relationship between the maximum load and the annual electricity consumption [18]:

$$P = C_1 E + C_2 \sqrt{E}, \tag{39.12}$$

where P is the annual maximum load, E is the annual electricity consumption, and C_1 and C_2 are model parameters. By using regression analysis with historical data, we can get the values of C_1 and C_2.

With the load data and the annual electricity consumption data of Fujian province from 2000–2009, according to the formula (39.12), we can get the relational model between maximum load and annul electricity consumption, as follows:

$$P = 17.4951E - 47.4899\sqrt{E}, \tag{39.13}$$

where the unit of P is MW and the unit of E is 100 million kWh. Put annual electricity consumption 499.99 billion kWh into the model. The final saturated load value is 84,739 MW, which has only a difference of 0.3 % compared with the predictions of the SD method.

Compared with the SD method, although the PCEC method is easier to operate, its prediction accuracy is not precise enough because, on the one hand, it only uses historical load and consumption data to predict and, on the other hand, forecasting maximum load utilization hours is not easy. On the contrary, the SD method has the advantages of being multidisciplinary. The established model is scalable, which can be adjusted according to different forecasting objects or expanded into a more detailed system. We also can increase the accuracy of the forecast during the process as we have accumulated more experience of the parameter assignment.

39.6 Conclusion

In this chapter, we first establish the SEF model. And then the SLF model is built on the basis of the SEF model with the method of SD. The prediction results are satisfactory, and the physical significance of the parameters is clear. What is more, the model combines the advantages of being multidisciplinary, making it more scientific and comprehensive than the previous load forecasting method. With the progressive development of the power system and electricity market, the SD method will increasingly demonstrate its superiority in power system applications.

Acknowledgments The research was financially supported by the Natural Science Foundation of Jiangsu province (No. BK20130742).

References

1. Daneshi H, Daneshi A. Real time load forecast in power system, in 3rd International Conference on Deregulation and Restructuring and Power Technologies, IEEE; 2008. p. 689–95.
2. Hyndman RJ, Fan S. Density forecasting for long-term peak electricity demand. IEEE Trans Power Syst. 2010;25(2):1142–53.
3. Dong GL, Byong WL, Soon HC. Genetic programming model for long-term forecasting of electric power demand. Electr Power Syst Res. 1997;40(1):17–22.
4. Cunkas M, Altun AA. Long term electricity demand forecasting in Turkey using artificial neural networks. Energy Sour Part B. 2010;5(3):279–89.
5. Kandil MS, El-Debeiky SM, Hasanien NE. Long-term load forecasting for fast developing utility using a knowledge-based expert system. IEEE Trans Power Syst. 2002;17(2):491–6.
6. Towill DR. System dynamics—background, methodology, and applications. Part 1: background and methodology. Comput Control Eng J. 1993;4(5):201–8.
7. Towill DR. System dynamics-background, methodology, and applications part 2: applications. Comput Control Eng J. 1993;4(6):261–8.
8. Hongming Y, Gaojie W, Lixing Z, Renjun Z. A study of power market dynamics based on system dynamics modeling. In 2006 International Conference on Power System Technology, IEEE/PES, 2007. p. 1–6.
9. Pappas SS, Ekonomou L, Karampelas P, Karamousantas DC, Katsikas SK, Chatzarakis GE, Skafidas PD. Electricity demand load forecasting of the Hellenic power system using an ARMA model. Electr Power Syst Res. 2010;80(3)256–64.
10. Hongbo Z, Nian Z. Forecast of energy demand in the next decade. Energy Procedia. 2011;5:2536–39.
11. Zhou P, Ang BW, Poh KL. A trigonometric grey prediction approach to forecasting electricity demand. Energy. 2006;31(14):2503–11.
12. Piepel GF. The statistical analysis of compositional data. Technometrics. 1988;30(1):120–21.
13. Wang H, Liu Q. Compositional data predicting model and its application in industrial structure trend in China. Global Manage Rev. 2002;(5):27–9. (In Chinese).
14. Kermanshahi B, Iwamiya H. Up to year 2020 load forecasting using neural nets. Int J Electr Power Energy Syst. 2002;24(9):789–97.
15. Kong H, Hui H. Forecast on Hebei electronic information industry based on modern industrial system. In 2nd International Conference on Information Science and Engineering, IEEE, 2010. p. 1212–15.
16. Chui F, Elkamel A, Surit R, Croiset E, Douglas PL. Long-term electricity demand forecasting for power system planning using economic, demographic and climatic variables. Eur J Ind Eng. 2009;3(3):277–304.
17. Yao Y, Lian Z, Liu S, Hou Z. Hourly cooling load prediction by a combined forecasting model based on analytic hierarchy process. Int J Therm Sci. 2004;43(11):1107–18.
18. Wei W, Ting-ting F. The application of per-person electricity consumption method in saturation load forecasting. Power Demand Side Manage. 2012;14(1):21–3 (in Chinese).

Chapter 40
Multi-Scale Fault Frequency Extraction Method Based on EEMD for Slewing Bearing Fault Diagnosis

Jie Yang, Jie Chen, Rongjing Hong and Hua Wang

Abstract In view of the large low-speed slewing bearing, the vibration signals are always very weak and overwhelmed by other strong noise, which makes fault feature extraction from the signals very difficult. In order to solve this problem, a denoising method based on multi-scale principal component analysis (MSPCA) and the ensemble empirical mode decomposition (EEMD) is proposed with a new intrinsic mode functions (IMFs) selection strategy. After that, the vibration signal is reconstructed by the selected IMFs. Finally, a method of multi-scale fault frequency extraction of slewing bearing based on EEMD is applied to denoise the vibration signals. The application of this method is demonstrated with laboratory accelerated slewing bearing life test data. Results show that EEMD-MSPCA is more effective in multi-scale fault frequency extraction of low-speed slewing bearing.

Keywords Slewing bearing · Ensemble empirical mode decomposition (EEMD) · Principal component analysis (PCA) · Denoising · Fault diagnosis

40.1 Introduction

Slewing bearing is widely used in wind turbines, engineering machineries, coal mine machineries, marine platform, military equipment, and other large rotating machinery. As the key component of large rotating machinery. when it has failure, it will cause great economic loss and sometimes may endanger the security of the operator. Therefore, strengthen slewing bearing diagnostic technique of research work has significant to prevent potential failures of the slewing bearing.

In recent years, some scholars have done some research on the fault diagnosis of slewing bearing. Liu et al. [1] put forward a brief review about the fault diagnosis of slewing bearing based on vibration signal, temperature signal, friction torque signal, acoustic emission, and stress wave. Wahyu et al. [2] used the adaptive multi-scale decomposition characteristics of the ensemble empirical mode decomposition

J. Yang (✉) · J. Chen · R. Hong · H. Wang
School of Mechanical and Power Engineering, Nanjing Tech University, 211816 Nanjing, China
e-mail: yangjie@njtech.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_40

(EEMD) for analyzing the vibration signals collected from lab slewing bearing subject to an accelerated life test and the real-case data from a sheet metal company. Matej et al. [3] presented a new method in terms of fault diagnosis of slewing bearing based on EEMD and principal component analysis (PCA), and its validity is shown by a simulation fault slewing bearing experiment. After that, Matej et al. [4] considered the nonlinear characteristics of slewing bearing vibration and then improved the combined method of a previous publication in PCA parts with kernel PCA. Recently, Wahyu [5] presented a novel application of circular domain features such as circular mean, circular variance, circular kurtosis, and circular skewness calculation-based condition monitoring method for low-speed slewing bearing.

Through the research on the slewing bearing fault diagnosis above, some problems have been found, that is most of slewing bearing failure was artificially simulated and the effectiveness of the presented method was proved by simulation experiments; therefore, we performed a slewing bearing accelerated life test to obtain vibration signal of slewing bearing from its normal condition to fault and then to final failure. This study is concerned with finding a reliable method for low-speed slewing bearing nature defects fault diagnosis.

40.2 Basic Theory of PCA and EEMD

PCA is a feature extraction method which transforms a high-dimensional into a lower-dimensional subspace which contains most of the original data information. This prepares the ground for multi-scale principal component analysis (MSPCA) denoising which is explained in Sect. 3. For example, one matrix X can be decomposed into a score matrix T and a loading matrix P using PCA, its calculating formula is as follows:

$$X = TP^T + E = \sum_{m=1}^{k} t_j p_j^T + E \quad (k < m), \tag{40.1}$$

where $\boldsymbol{T}_{n\times k} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_k]$, $\boldsymbol{P}_{m\times k} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_k]$, E is the residual matrix and k is the number of selected principal components.

When a PCA model is built and a new data sample x is to be tested for fault detection, its calculating formula is as follows:

$$\boldsymbol{x} = \hat{\boldsymbol{x}} + \bar{\boldsymbol{x}} = \boldsymbol{P}\boldsymbol{P}^T\boldsymbol{x} + (\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^T)\boldsymbol{x}, \tag{40.2}$$

where $\hat{\boldsymbol{x}} = \boldsymbol{P}\boldsymbol{P}^T\boldsymbol{x}$ is the projection on the principal component subspace, and $\bar{\boldsymbol{x}} = (\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^T)\boldsymbol{x}$ is the projection on the residual subspace.

In this chapter, we use Squared prediction error (SPE) statistic to detect the fault sample. For a new scale measurement sample $\boldsymbol{X} \in \boldsymbol{R}^{n\times m}$, the SPE statistic is respectively computed as follows:

$$SPE = \left\| \bar{\boldsymbol{x}} \right\|^2 = \left\| (\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^T)\boldsymbol{x} \right\|^2. \tag{40.3}$$

For the fault detection purpose, we can compare the value of the SPE statistic with the corresponding threshold limit SPE_{lim}. The fault detection logic is if $SPE \leq SPE_{lim}$ it is fault-free, otherwise faulty [6, 7].

In the fault diagnosis of machinery, the method of empirical mode decomposition (EMD) has developed more extensive researches and applications in recent years [8, 9]. For the limitation of this thesis, the algorithm of EMD is not specifically discussed; however, the nonlinear components in nonstationary signals may result in a mode-mixing effect using the EMD method. Therefore, the EEMD algorithm was developed from the original EMD method by Wu and Huang [9], which added white noise into EMD procedures to restrain the mode-mixing effect. The EEMD decomposition algorithm can be shortly given as follows. First, initialize the number of ensemble M and the amplitude of the added white noise β. After that, follow the steps given below:

a. Add a white noise series with magnitude β to the original signal $x(t)$ to generate a new signal $y_m(t)$.

$$y_m(t) = x(t) + \beta n_m(t), \tag{40.4}$$

where ym(t) represents the noise-added signal of the mth trial, and nm(t) indicates the mth added white noise series.

b. Decompose the signal $y_m(t)$ into k intrinsic mode functions (IMFs) $C_{i,m}$ ($i = 1, 2, \ldots, k$) using the EMD method. $C_{i,m}$ represents the ith IMF of the mth decomposition, and k is the number of IMFs.
c. If $m < M$, then go to step (a) with $m = m + 1$. Repeat steps (a) and (b) M times with a different white noise series each time to obtain an ensemble of IMFs.
d. Calculate the overall ensemble mean of the corresponding IMFs of the decomposition as the final result:

$$C_i(t) = \frac{1}{M}\sum_{m=1}^{M} c_i(t) \quad (i = 1, 2, \ldots, k, m = 1, 2, \ldots M). \tag{40.5}$$

40.3 Multi-Scale Denoising Method Based on EEMD and PCA

The proposed multi-scale denoising method has been improved by the wavelet multi-scale decomposition algorithm [6]. The EEMD technique was used instead of wavelet analysis (WT) because the wavelet transform is a nonadaptive decomposition algorithm, which has the shortcoming that its analysis results depend on the choice of the wavelet base function. However, the full data-driven EEMD does not suffer from this limitation due to its adaptive empirical nature. Figure 40.1 shows the multi-scale denoising and fault frequency extraction mode.

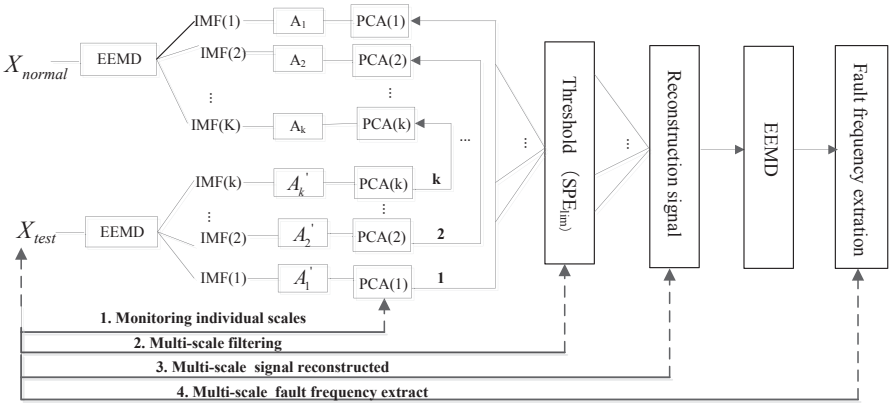


Fig. 40.1 EEMD multi-scale fault frequency extraction mode

The proposed method of multi-scale denoising technique requires two steps. In the first step, the normal vibration signal is decomposed by EEMD into k different timescale IMFs, then a normal MSPCA mode is constructed by these single IMFs, and the normal statistical limits SPElim of each timescale PCA model are calculated. In the second step, the newly measured vibration signal is also decomposed into k IMFs using the EEMD method; then the newly decomposed IMFs component is projected onto the corresponding normal PCA models constructed in the first step. Lastly, according to the statistical limits SPElim it is determined whether each timescale mode follows normal operation or not, and the certain timescale (IMFs) which has the failure existed to reconstruct the signal is selected.

40.4 Analysis of Real-Case Slewing Bearing Vibration Signal

In order to further verify the efficiency and performance of the proposed method for the real-case slewing bearing vibration signal denoising and fault diagnosis, the proposed method was applied to the slewing bearing accelerated life experimental data and the results, analysis, and description are presented below.

40.4.1 Slewing Bearing Experiment Bench

The experiment bench is an independent design and developed by our team to simulate the actual operation conditions of large-size bearing which can be used in a variety of applications experiment. Figure 40.2 shows the slewing bearing test bench which can bear axial load, radial load, and overturning torque. In this experiment, the test bench operated in one direction of rotation at a speed of 4 rpm. In

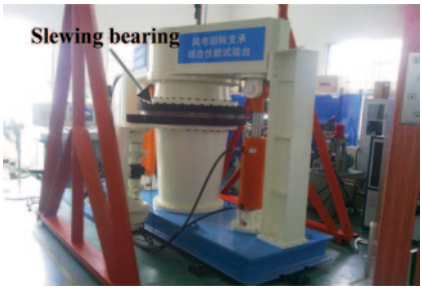


Fig. 40.2 Slewing bearing test bench

Table 40.1 Slewing bearing fault frequencies

Defect mode	Fault frequencies/Hz
Outer ring (BPFO)	3.18
Inner ring (BPFI)	3.28
Rolling element (BSF)	2.21
Cage train	0.49

this slewing bearing accelerated life test, four accelerometer sensors were placed on the slewing bearing at 90° to each other. The slewing bearing fault frequencies are shown in Table 40.1.

40.4.2 Analysis of Experimental Data

In order to ensure the extracted vibration data, which provides sufficient fault information, we must extract enough vibration data to analyze it. Referring to the fault frequencies presented in Table 40.1, the lowest frequency was 0.49 Hz. After simple calculation, 0.49 Hz is 2 s, we should extract at least 2-s-vibration data that can ensure to represent all fault frequencies; therefore, in this chapter, we extract 3-s-vibration data which are enough to represent all fault information and reduce the EEMD algorithm computation time.

The slewing bearing accelerated life test began on May 16, 2013. In order to accelerate the slewing bearing defect, the slewing bearing was continuously running with full load. After continuous operation of slewing bearing for 12 days (up to May 28), we were aware that failure could occur in slewing bearing on the basis that vibration and noise increase. Therefore, according to the EEMD multi-scale failure frequency extracted mode, two vibration data groups were analyzed: the vibration signal at midnight on May 16 as the normal sample (Fig. 40.3) and the vibration signal at midnight on May 28 as the fault sample (Fig. 40.4). Nevertheless, when we used the EEMD multi-scale fault frequency extraction mode, two parameters should be determined in EEMD method in order to obtain an optimum decompo-

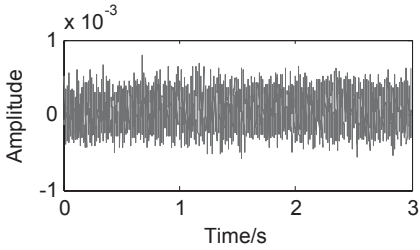


Fig. 40.3 Vibration signal on May 16

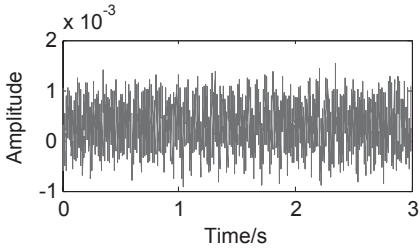


Fig. 40.4 Vibration signal on May 28

sition result. In this chapter, the white noise magnitude β of 0.02 is used and ensemble number E of 100 is selected. The vibration signal on May 28 after filtering is shown in Fig. 40.5 on the basis of the steps described in Sect. 3.1. Then, the fault frequencies of the filtering signal were extracted by EEMD as shown in Table 40.2. According to the results in Table 40.2, the frequencies 2.33 Hz and 3 Hz are close to the corresponding characteristic defect frequency of the rolling element and outer ring (BPFO) in Table 40.1, so we infer that the fault occurs in the rolling element and the outer ring. Finally, slewing bearing was removed to inspect each component, and it was found that some rolling elements have been damaged, and there appeared a little bit of corrosion in the outer ring. The damaged rolling element and outer ring were shown in Figs. 40.8 and 40.9. Therefore, the fault diagnosis method proposed in this chapter was applied to the slewing bearing accelerated life experiment, and the results verify its validity and feasibility (Figs. 40.6 and 40.7).

40.5 Conclusion

Based on EEMD and principal component analysis, this chapter proposes a method in combination of the multi-scale denoising and the multi-scale fault frequency extraction of slewing bearing fault diagnosis. In the presented method, the method of multi-scale denoising with EEMD is performed to preprocess slewing bearing vibration signals. Through the slewing bearing accelerated life experiment, vibration signal analysis results demonstrate that the proposed method can reliably recognize the faults.

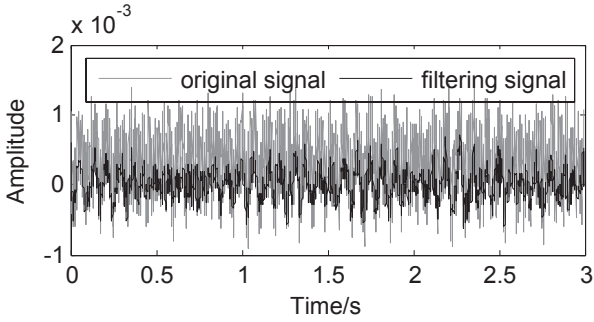


Fig. 40.5 Result of multi-scale denoising

Table 40.2 EEMD multi-scale fault frequency extraction results (May 28)

IMF	Frequency/Hz	IMF	Frequency/Hz
1	759.7	7	5.333
2	240.7	8	3
3	145.3	9	2.333
4	68	10	1.667
5	26.67	11	0.667
6	14	12	0.333

IMF intrinsic mode function

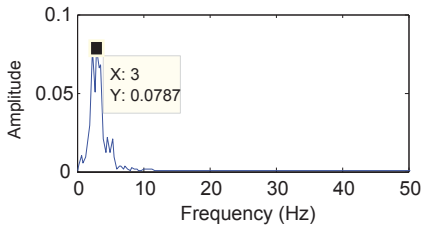


Fig. 40.6 Outer ring fault frequency

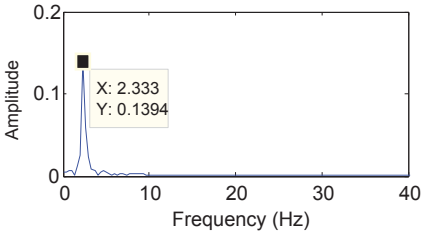


Fig. 40.7 Rolling element fault frequency

Fig. 40.8 Outer race failure

Fig. 40.9 Rolling element failure

References

1. Liu Z, Chen J. Monitoring and diagnosis technique on slewing bearing. Mod Manuf Eng. 2011;156–157(11):127–31 (In Chinese).
2. Caesarendra W, Kosasih P, Tieu KA, et al. Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition. J Mech Sci Technol. 2013;27(08):2253–62.
3. Zvokelj M, Zupan S, Prebil I. Multivariate and multiscale monitoring of large-size low-speed bearings using ensemble empirical mode decomposition method combined with principal component analysis. Mech Syst Signal Process. 2010;24(04):1049–67.
4. Zvokelj M, Zupan S, Prebil I. Non-linear multivariate and multiscale monitoring and signal denoising strategy using kernel principal component analysis combined with ensemble empirical mode decomposition method. Mech Syst Signal Process. 2011;25(07):2631–53.
5. Caesarendraa W, Kosasih B, Tieu KA, et al. Circular domain features based condition monitoring for low speed slewing bearing. Mech Syst Signal Process. 2014;45(01):114–38.
6. Misraa M, Yue H. Multivariate process monitoring and fault diagnosis by multi-scale PCA. Comput Chem Eng. 2002;26(09):1281–93.
7. Ding S, Zhang P, Ding E, et al. On the application of PCA technique to fault diagnosis. Tsinghua Sci Technol. 2010;15(02):138–44.
8. Lei Y, He ZJ, Zi YY. Application of the EEMD method to rotor fault diagnosis of rotating machinery. Mech Syst Signal Process. 2009;23(04):1237–338.
9. Zhang Y, Zuo H F, Bai F Classification of fault location and performance degradation of a roller bearing. Measurement. 2013;46(03):1178–89.

Part IV
Transmission and Distribution

Chapter 41
Comparison of Form-Finding Methods for Transmission Lines Based on ANSYS

Ying Jin, Changsheng Liu, Yunyun Xie and Linghao Zhang

Abstract ANSYS is the most popular finite element analysis software all over the world. Although there are more than one kind of form-finding method for transmission lines, study on the comparison of these methods has not been carried out before. Firstly, this chapter will summarize the theory and the algorithm flow of three common form-finding methods for the initial state of transmission lines based on ANSYS, specifically, they are the "V"-shape method, the little elasticity modulus method and the direct iteration method; besides, this chapter believes that these three methods have distinct accuracies in different calculation conditions. With a series of ANSYS simulation examples, the errors between simulation results and theoretical values are expressed by scatter charts. Furthermore, this chapter discusses the simulation results by comparing the accuracy of these three form-finding methods.

Keywords Comparison · Transmission lines · ANSYS · Form finding

41.1 Introduction

Ideally, the cable-structure does not possess the compression performance and bending resistance, and the material of cable-structure is supposed to conform to Hooke's law [1–4]. After being applied with prestress, the cable-structure becomes rigid with the affordability of external loads [2, 5]. Actually, the initial form-finding methods for conductor are the foundation of mechanical analysis of tower-line coupling system [3–5]; therefore, the precision of form-finding results will exert great influence on the subsequent dynamic analysis.

There are three common form-finding methods for transmission lines based on ANSYS, namely, the "V"-shape method, the little elasticity modulus method, and the direct iteration method. Lu et al. propose the "V"-shape method, which is characterized by the initial "V"-shape geometric model [6, 7]. Jia et al. propose the little elasticity modulus method, which needs adjusting the elasticity modulus value to meet the demands of iteration stopping [8–10]. Chen et al. propose the direct

Y. Xie (✉) · Y. Jin · C. Liu · L. Zhang
School of Automation, Nanjing University of Science and Technology, 210094 Nanjing, China
e-mail: xyy_njust@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_41

iteration method, in which all the parameters are set in accordance with practical values [11–14].

These form-finding methods have many differences in theory and precision; therefore, this chapter mainly discusses and compares these three methods based on a group of ANSYS examples, with some conclusions about the accuracy of these three form-finding methods for transmission conductor.

41.2 Form-Finding Methods for Transmission Lines Based on ANSYS

41.2.1 "V"-shape Method

The geometric model of "V"-shape method is a "V"-shape model with three key points: two suspension points and the lowest point [6, 77]. The significant advantage of "V"-shape method is simple to operate, and the algorithm flow of "V"-shape method is shown in Fig. 41.1.

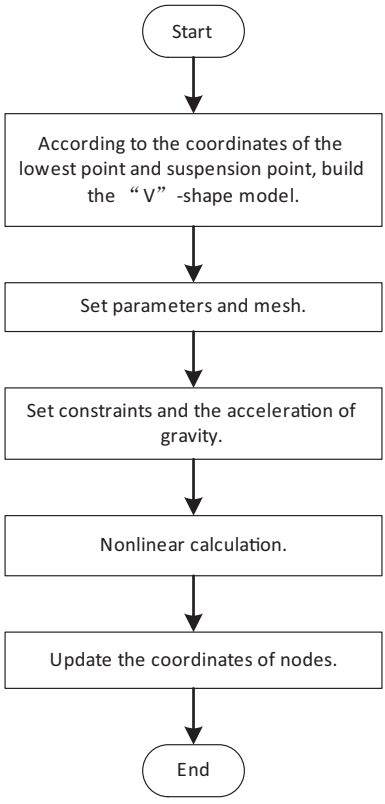


Fig. 41.1 Algorithm flow of the "V"-shape method

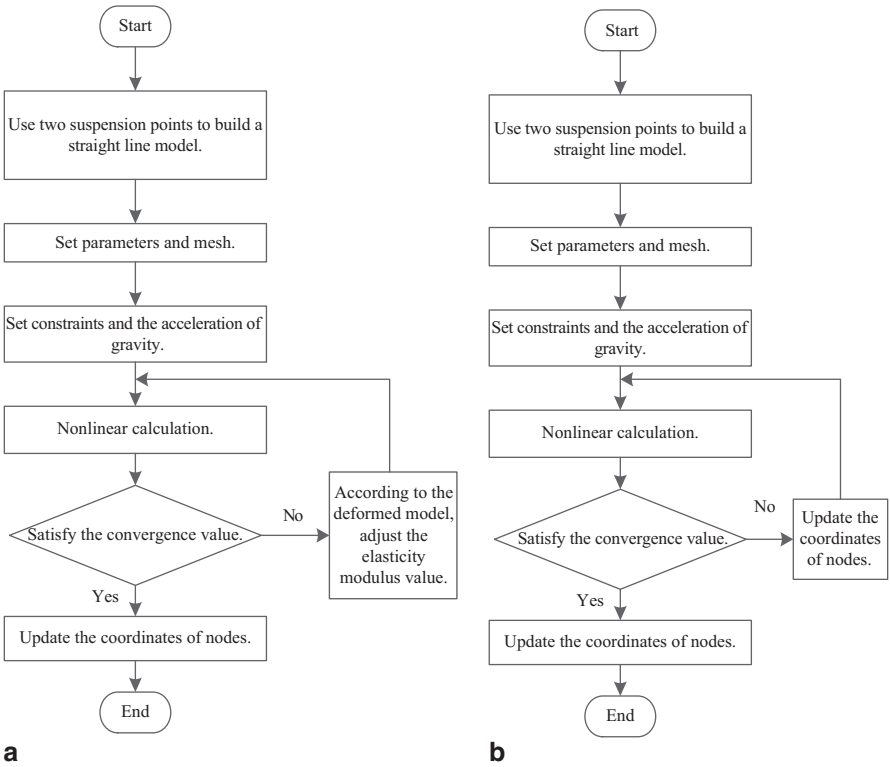


Fig. 41.2 Algorithm flow of little elasticity modulus method and direct iteration method

41.2.2 Little Elasticity Modulus Method

The core of little elasticity modulus method finds the initial form of transmission conductor in combination with the gravity load and smaller elasticity modulus value [8–10]. The operation of little elasticity modulus method is complicated, which is shown in Fig. 41.2a.

41.2.3 Direct Iteration Method

In the direct iteration method, the length increase of catenary after form finding is simply considered as the result of gravity [11–14], and the iteration cannot stop until its results meet the requirement of accuracy. The algorithm flow of this method is shown in Fig. 41.2b.

41.3 Examples

41.3.1 Design of Examples

Considering factors of span distance and height difference, this chapter has designed three kinds of calculation models. Among them, only one model has height difference with other models having no height difference. The calculation conditions of each model are described in Table 41.1.

41.3.2 Results

This chapter analyzes the precision of two types of data: the length of catenary and the stress of the suspension point, and the corresponding theoretical values of these data can be calculated by the catenary equation [15, 16]; furthermore, the scatter charts of errors between theoretical values and calculation results are shown in Figs. 41.3, 41.4, 41.5.

Table. 41.1 Calculation conditions

Calculation conditions		Descriptions of calculation conditions
Model I	S-1-1	No height difference; span is 50 m
	S-1-2	No height difference; span is 300 m
	S-1-3	No height difference; span is 450 m
Model II	S-2-1	height difference is 2 m; span is 200 m
	S-2-2	height difference is 6 m; span is 200 m
	S-2-3	height difference is 10 m; span is 200 m
Model III	S-3-1	height difference is 30 m; span is 200 m
	S-3-2	height difference is 40 m; span is 200 m
	S-3-3	height difference is 50 m; span is 200 m

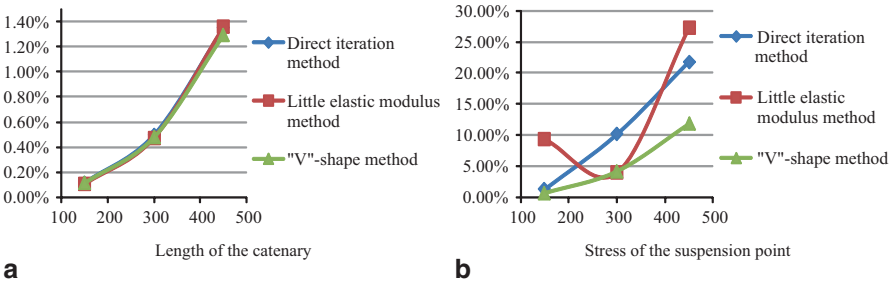


Fig. 41.3 Scatter chart of model I

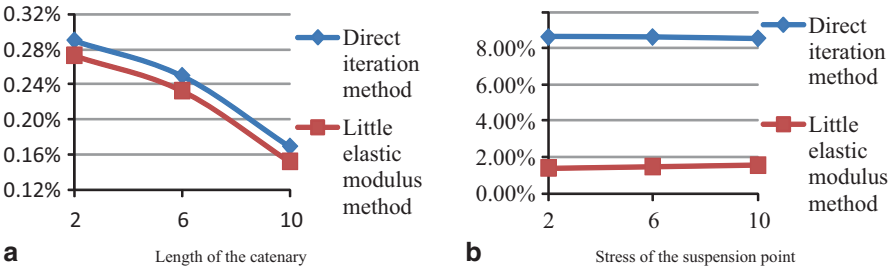


Fig. 41.4 Scatter chart of model II

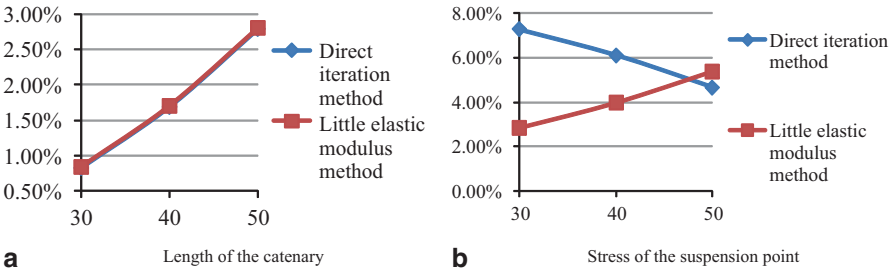


Fig. 41.5 Scatter chart of model III

41.4 Discussion

Results indicate that in different calculation conditions, the accuracy of each form-finding method for transmission lines changes a lot.

The "V"-shape method cannot be used in the situation where the height difference is quite large because the principle of this method is invalid in such situation. In model I, there is no height difference, and the accuracy of "V"-shape method decreases with the increment of span distance. Actually, the shorter the span is, the harder transmission lines will deform, which indicates if the "V"-shape method is used in the situation where the span is short, the error of its simulation results is small.

The accuracy of little elasticity modulus method differs extremely because of variation of calculation condition. Firstly, there is no unified standard for setting the initial elasticity modulus; therefore, it greatly depends on personal experience, which means the large error might be produced; besides, the criterion of iteration stopping is not unified.

Similarly, the accuracy of direct iteration method varies a lot in the distinct calculation condition. For example, the stress-error of direct iteration method in model I is rising steeply with the increase of span; at the same time, it is falling down in model III. Specifically, the main reason that the accuracy of direct iteration method changes with the calculation condition is the lack of standard and unified convergence value.

Unfortunately, we do not figure out the selection strategy of these form-finding methods, which needs further study. Beyond doubt, with the help of selection strategy, we can find the best form-finding method for transmission lines quickly according to the given conditions.

41.5 Conclusion

Generally speaking, we believe that the accuracy of these three form-finding methods for transmission lines changes a lot with the variation of calculation condition; moreover, further study on selection strategy of form-finding method for the initial state of transmission lines should be carried out.

References

1. Li B, Gong PH. Self-study manual ANSYS14 of finite element analysis. Beijing: Posts and Telecom Press; 2013. p. 239–50 (In Chinese).
2. Wang XM. Numerical analysis of ANSYS engineering structure. Beijing: China Communications Press; 2007. p. 196–210 (In Chinese).
3. Wang Y. Analysis and realization of form-finding of cable-membrane structure by finite element method. J Liaoning Tech Univ. 2011;24(4):104–7 (In Chinese).
4. Xia ZC, Li L, Liang ZP, et al. Dynamic response of transmission tower with ruptured wires. Vibration Shock. 2007;26(11):45–9 (In Chinese).
5. Veenendaal D, Block P. Overview and comparison of structural form-finding methods for general networks. Solids Struct. 2012;49:3741–53.
6. Lu YB. Analysis of ice-shedding and wire-breaking in transmission line system. Hangzhou: Zhejiang University; 2010 (In Chinese).
7. Lu YB, Lou WJ, Li HL. Process analysis of non-uniform ice shedding of transmission lines. Vibration Shock. 2010;29(9):47–50 (In Chinese).
8. Jia YZ, Liu RP, Li ZQ. Original state for overhead icing transmission line. Water Res Power. 2011;29(1):148–50 (In Chinese).
9. Meng SM, Shan LP. Form-finding in dynamics analysis of transmission line. Adv Power Sys Hydroelectr Eng. 2009;25(10):43–7 (In Chinese).
10. Yang Q, Li CM. Form-finding and simulation of cable structure in ANSYS. Info Technol Civil Eng Archit. 2010;2(4):61–5 (In Chinese).
11. Chen SL, Cao MG. Analysis of form-finding of overhead transmission lines and grounding line based on the application of ANSYS. Jiangxi Electric Power. 2007;31(6):7–15 (In Chinese).
12. Kong W, Zhu MW, Fu H. Research on form-finding method for transmission lines in ANSYS. J Northeast Electr Power Univ. 2012;31(5/6):64–7 (In Chinese).
13. Zhang WH. Numerical simulation and experimental research on vibration of iced overhead conductors. Beijing: North China Electric Power University; 2012 (In Chinese).
14. Zhang WH, Yu JB. Form-finding analysis of overhead transmission line based on ANSYS. Electr Power Constr. 2012;33(2):32–5 (In Chinese).
15. Shao TX. Mechanical calculation of overhead transmission lines. 2nd edn. Beijing: China Electric Power Press; 2003. p. 181–90 (In Chinese).
16. Zhang DS. Design manual of high voltage transmission lines in electric power engineering. 2nd edn. Beijing: China Electric Power Press; 2002. p. 177–98 (In Chinese).

Chapter 42
Simulation Analyses and Realization of Lightning Current Monitoring System on Transmission Lines

Jian Wei, Lu Qu, Hailiang Lu, Yu Wang, Lei Lan and Xishan Wen

Abstract Based on the production and operation situation of the State Grid Corporation of China in recent years, the lightning faults have taken of the majority of the transmission and distribution equipment faults. In order to prevent the lightning damages and control the lightning faults of the transmission lines, it is necessary to study the impacts of the lightning strike on transmission lines. In the course, the primary work is to accurately judge the location and form of the lightning strike. Upon simulation of the current waveform on transmission lines subject to different forms of lightning attacks, the features of lightning current on the insulators and grounding lines have been achieved; then, the judging methods of the polarity and the location of the lightning strike are studied. In order to monitor the waveform, amplitude, polarity and other parameters of lightning current in a real-time manner, a transmission line lightning monitoring system is developed. Based on the judging methods achieved by simulation and monitoring the waveform, the lightning fault location and the form of lightning can be confirmed. The monitoring result shows that this system can provide effective technical guarantee to take pertinent lightning protection measures.

Keywords Lightning current waveform · Real-time monitoring · Lightning parameter confirming

H. Lu (✉) · J. Wei · L. Qu · Y. Wang · L. Lan · X. Wen
College of Electrical Engineering, Wuhan University, 420072 Wuhan, China
e-mail: luhailiang@whu.edu.cn

J. Wei (✉)
e-mail: weijian6611@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_42

42.1 Introduction

The lightning striking transmission lines may bring about great harm to the transmission network. According to the classification of power network failure in China, the breakdown due to lightning striking is 40 % of the total amount or so in the area of high trip-out rate [1–2]. Obviously, the lightning brings huge economic loss. Ascertaining the lightning striking points and the lightning current polarity as well as collecting the lightning current waveform are of great significance for studying the lightning striking features and reducing the harm that the lightning brings. At present, in the research and analysis on trip-out accident of lightning, existing problems of great workload lie in finding the lightning trip-out fault and the difficulty in identifying the lightning tripping reasons; in this sense, it is necessary to carry out the research of online monitoring of lightning current parameters. The installation of lightning measuring devices on the transmission lines and real-time monitoring of the line lightning location, lightning current waveform, and amplitude can provide strong support for seeking location of the lightning line fault, confirming the standards of lightning current waveform and amplitude, and thus exert positive effect on the development of lightning protection technology [3–4].

From 1994 to 2004, Tokyo Electric Power Company directly measured the current waveforms of lightning strokes on 60 transmission towers [5], but they only measured the current on towers and paid no attention to the current on lines. So it could not confirm the strike point on lines. In this chapter, we will study the confirmation method of strike point.

42.2 Simulation Analyses

In order to get the current on lines under different conditions, simulation of the lightning striking lines under different conditions is needed. Advanced graphics international electromagnetic transient program ATP-EMTP is used for simulation in this research. EMTP is the acronym of electromagnetic transient program. As one of the most widespread procedures in high-voltage power system simulation, it focuses on the operating condition of the system rather than the individual switch details.

In the simulation model, the transmission line tower is regarded as the distributed parameter model. The wave impedance and velocity of propagation in the tower are used to represent the lightning wave propagation characteristics in the tower.

The operating voltage of simulation is set at 35 kV. At this voltage level, as short-circuit current is much less than the lightning current, it is not considered in this simulation. The monitoring nodes are as shown in Fig. 42.1. Lightning current waveform selects the 8/20 μs standard lightning waveform, whereas the influence of line arresters is not considered in the simulation.

Upon the establishment of the tower model, five current monitoring points are set on the three-phase insulators and lightning line. Then, the lightning striking point and the lightning parameter are set, and the current on monitoring points is

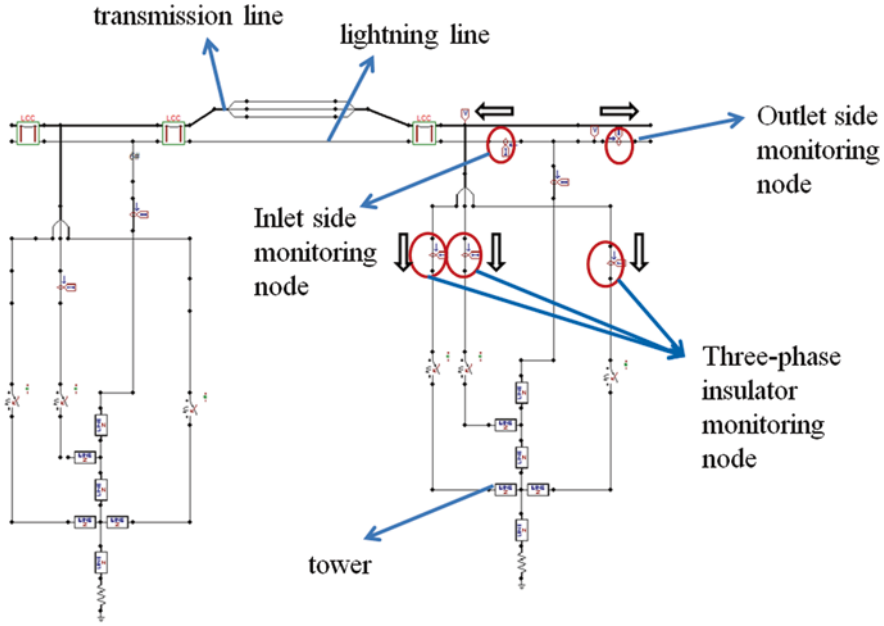


Fig. 42.1 Monitoring node position and the positive direction in simulation. (The *black hollow arrow* indicates the specified current positive direction)

recorded when the lightning striking point and the lightning form are different. Conclusions are obtained by analyzing the simulation results.

This article simulates the situation that positive polarity or negative polarity lightning strikes the tower top, the lightning line at span middle of tower and the transmission conductor. Each situation is analyzed in two cases: insulator flashover and not flashover.

While the lightning current amplitude is high, insulator flashover occurs, and a part of lightning current will flow into transmission conductors through insulators. If the amplitude of lightning current is not huge enough to cause the insulator flashover, the majority of the lightning current will flow into the earth through the tower with rest of lightning current flowing away through ground wire. The lightning current on insulators and lightning conductor can be used to judge the striking location and the polarity of lightning with the judgment shown in Table 42.1.

For example, if the measured current on lightning line inlet side of tower is positive and of large amplitude, the current on lightning line outlet side of tower is negative and of small amplitude, and the current on insulators is negative and of small amplitude; as a result, the lightning is positive and strikes on the inlet side of transmission conductor, and the flashover occurs.

The simulation results show that when the insulator flashover, the lightning strikes the lightning line, the current amplitude on the lightning line is proportional to the current amplitude of lightning, and the current amplitude on tower inlet side and the outlet side is approximately equal. The situation that the lightning strikes

Table 42.1 Judgment method of lightning polarity and striking location

Current on lightning line inlet	Current on lightning line outlet	Current on insulators	Conclusion of lightning
P, L	P, L	-,-	P, ST, NF
P, L	N, M	-,-	P, SIL NF
N, M	P, L	-,-	P, SOT, NF
N, L	N, L	-,-	N, ST, NF
N, L	P, M	-,-	N, SIL NF
P, M	N, L	-,-	N, SOT, NF
P, L	P, L	N, S	P, ST, F
P, L	N, M	N, M	P, SIL F
N, M	P, L	N, M	P, SOL,F
N, S	N, M	N, L	P, SIT, F
N, M	N, S	N, L	P, SOT, F
N, L	N, L	P, S	N, ST, F
N, L	P, M	P, M	N, SIL F
P, M	N, L	P, M	N, SOL,F
P, S	P, M	P, L	N, SIT, F
P, M	P, S	P, L	N, SOT, F

P positive, *N* negative, *L* large current amplitude, *S* small current amplitude, *M* medium current amplitude, *ST* strike on the top of tower, *SIL* strike on inlet side of lightning line, *SOL* strike on outlet side of lightning line, *SIT* strike on inlet side of transmission conductor, *SOT* strike on outlet side of transmission conductor, *NF* no flashover, *F* flashover, - no current

the transmission conductor without insulator flashover is rare, and it is difficult to detect.

In case of lightning flashover, when the lightning strikes the tower top, compared with no flashover, a part of lightning current flows to the transmission conductor through the insulator, and the induction voltage of the lightning current on the transmission conductor may affect the current on lightning line, and current amplitude difference of inlet side and the outlet side becomes larger. When the lightning strikes the lightning line at span middle of tower, the current amplitude at inlet side is larger, and the current amplitude at outlet side is small; most current flows into ground through the tower, and little current will flow into the transmission conductor by the back striking. When the lightning strikes the transmission conductor, the current amplitude of the nearer side to strike point of lightning line is very small, and it is hard for the monitoring device to monitor. So, it can be judged by the current on the insulator and the other side of lightning line. The theory can be used to judge the striking location and the polarity of lightning.

42.3 System Realizations

On this basis, the online monitoring system which has five channels has been made. As shown in Fig. 42.2, the lightning current monitoring system is composed of the monitoring node and the monitor terminal, where the monitoring node is composed of lightning current sensor, signal acquisition unit, wireless communication module and power-supply module. It can monitor the lightning current on the insulators and the lightning conductors at the same time. When the lightning strikes the lightning conductors or the transmission conductors, the lightning current sensor can get the induced voltage, and then the voltage signal is transmitted to the monitoring node through coaxial cable. The monitoring node will record the lightning current wave and the lightning strike time, and then store the monitoring result. The monitoring result will be transmitted to the monitor terminal by wireless communication network.

The monitor terminal is composed of industrial personal computer and GPRS module. The monitoring node is composed of high-speed analog-to-digital converter, MSP430 signal processing chip, FPGA and wireless communication module [6, 7]. The system structure is shown in Fig. 42.3.

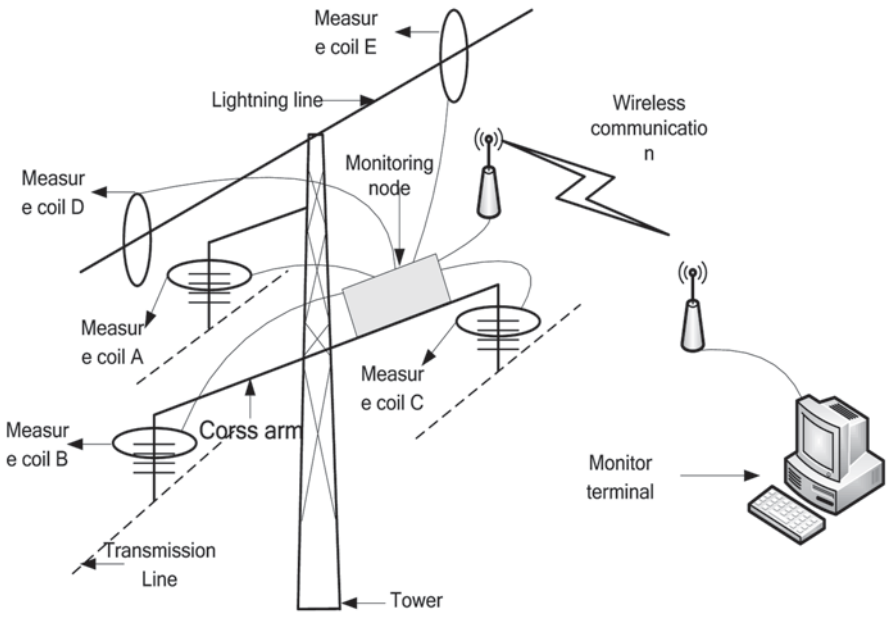


Fig. 42.2 Structure of the monitoring system

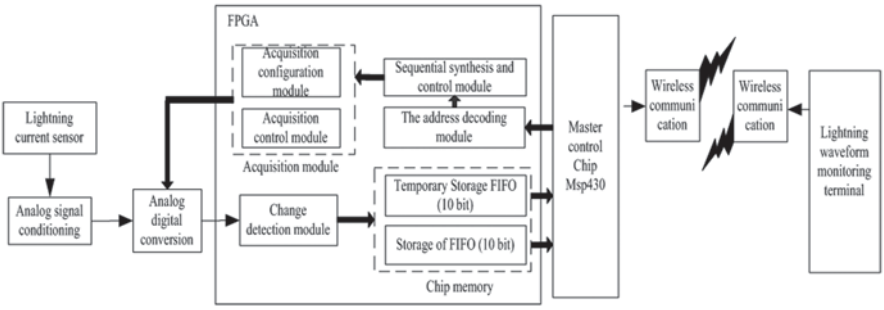


Fig. 42.3 Structure of monitoring node

The performance of sensor has a direct influence on the monitoring node. Rogowski coil can accurately measure the amplitude, the polarity and the waveform of lightning current. Its construction is simple and features very good linearity. Rogowski coil is chosen as the sensor. External-integral coil has a wider band than self-integral, so it is more suitable for measuring the lightning current. The device in this chapter uses the self-integral Rogowski coil [8].

As the monitoring nodes work in field for a long term, the lightning current acquisition unit device should consume low power. The core of digital system consists of high-speed analog-to-digital converter, MSP430 signal processing chip, and FPGA. This system chooses EP1C12Q240 chip of Altera Company as the logic chip, and chooses MSP430F5438 microcontroller of TI Company as the control chip. The 10-bit parallel chip ADS822 of TI Company is used as the high-speed analog-to-digital converter. The sampling rate of this analog-to-digital converter can reach 40MS/s with the typical power consumption of 200 mW.

Considering that the monitoring node and the monitor terminal are usually at a distance of more than 10 km, GPRS is the best choice of communication [9].

Industrial computer is used as the monitor terminal and LABVIEW is used to write a control program. The program flow chart of the software system is shown in Fig. 42.4. When the monitor node starts run, the sampling points are initialized first and the configuration of interruption opens. When a lightning current flows through the coil, the lightning and the voltage signal that the lightning induces will reach a certain amplitude, and the MSP430 external interrupt responds accordingly and the MSP430 chip enters the lightning acquisition state. After the completion of the acquisition, MSP430 reads the data of FIFO, packs the data with a communication protocol, and then sends the data to the monitoring terminal according to the priority through the GPRS module. The monitoring terminal reads the serial port data. When receiving a correctly packed data, the monitoring terminal will convert the codes and calculate the waveform parameters. Then, the data will be eventually stored in the computer hard disk in an EXCEL file.

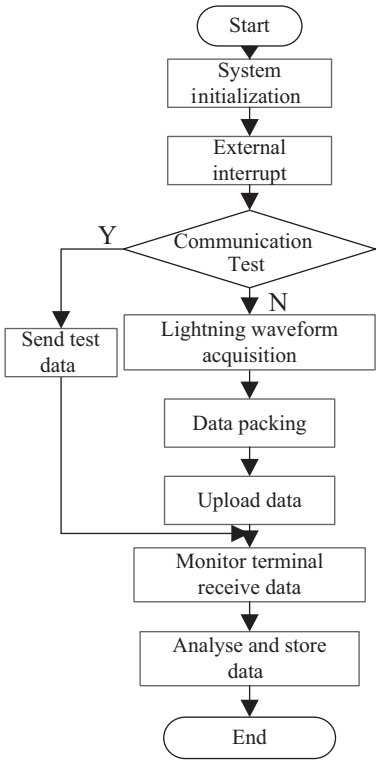


Fig. 42.4 Program flow chart of the software system

42.4 Performance Test and Installation

The monitoring system is tested in a laboratory, and a 8/20-μs impulse current generator is used. The performance test diagram is shown in Fig. 42.5, and the test result is shown in Fig. 42.6.

The monitoring system has been installed in Shanxi province in China and it is working normally, as shown in Fig. 42.7.

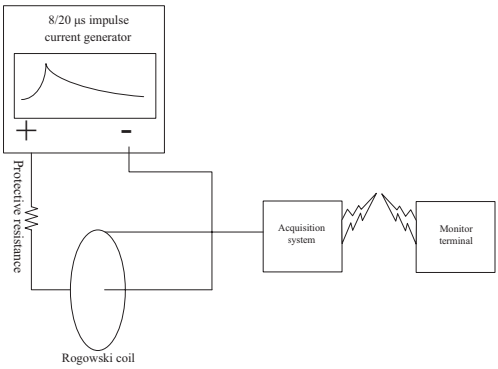


Fig. 42.5 Performance test diagram

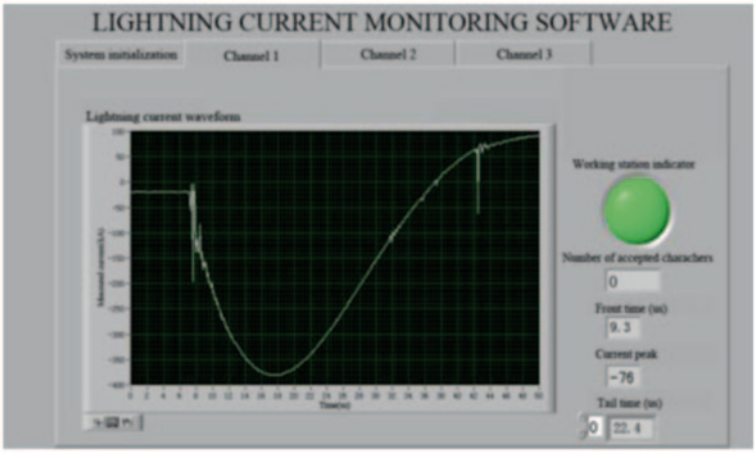


Fig. 42.6 Monitoring system testing

Fig. 42.7 Equipment installation

42.5 Conclusion

Upon a lot of simulations, the features of lightning current on the insulators and lightning lines have been obtained when the polarity and the location of the lightning strike are different; then, the judgment method of the polarity and the location of the lightning strike are gotten. A transmission line lightning current monitoring system based on high-speed analog-to-digital converter, MSP430 signal processing chip and FPGA is developed in order to monitor the current waveform, the amplitude, the polarity and other parameters in a real-time manner when the lightning strikes the ground lines, transmission lines or lightning towers, thus confirming the lightning fault location and the form of lightning. The installation of lightning measuring devices on transmission lines, real-time monitoring line lightning location, lightning current waveform and amplitude can greatly support the location seeking of lightning line fault, confirm the standards of lightning current waveform and amplitude, and thus exert a positive effect on the development of the lightning protection technology.

References

1. Shen H. research of impulse grounding resistance measurement equipment based on DSP. Nanchang: Nanchang University; 2006 (In Chinese).
2. Wang T, Zhou W, Yi H. Real time stroke current monitoring system for overhead transmission line. High Volt Eng. 2008;34(5);961–5 (In Chinese).
3. IEEE Working Group on Estimating Lightning Performance of Transmission Lines. A simplified method for estimating lightning performance of transmission lines. IEEE Trans. 1985;PAS-104(4):919–32.
4. IEEE Working Group on Estimating Lightning Performance of Transmission Lines. Estimating lightning performance of transmission lines ii-updates to analytical models. IEEE Trans. 1993;PWRD-8(3):1254–67.
5. Takami J. Shigemitsu Okabe. Observational results of lightning current on transmission towers. IEEE Trans Power Deliv. 2007;22(1):547–56 (January).
6. Zhang D. Development of multi-channel lightning current measurement device. Beijing: North China Electric Power University; 2009 (In Chinese)
7. Mochizuki ZE, et al. Observation of lightning by means of time-of-arrival type lightning location system. Electr Eng Japan (English Translation of Denkigakkai Ronbunshi). 1997;120(1):9–15.
8. Weibo Li. Study of sensor theory cantered on rogowski coil for heavy current measurement application. Wuhan: Huazhong University of Science and Technology; 2005 (In Chinese).
9. Zhou S, Gu C. Study on remote realtime overvoltage monitoring system based on embedded technology. High Volt Eng. 2004;12(1):53–5 (In Chinese).

Chapter 43
Analysis of Back-Flashover Lightning-Withstanding Level of 500-kV-Double-Circuit Transmission Lines

Shihong Wu, Xishan Wen, Rui Li and Yang Xu

Abstract This chapter analyzes the lightning performance of 500-kV-double-circuit transmission lines by the Alternative Transient Performance-Electromagnetic Transient Program (ATP-EMTP). With towers using the multi-wave impedance model and eliminating voltage reflection on the lines by a matching resistor, a simulation method has been established to evaluate the factors that may influence the back-flashover voltage such as tower height, tower footing resistance, and conductor phase arrangements. A statistical method is presented to calculate the line trip-out rates due to back-flashover. The results show that the impact of phase arrangements on the back-flashover trip-out rate of a single-circuit line is little, but significant on the double-circuit lines.

Keywords Double-circuit line · Back-flashover · Lightning performance · Trip-out rate · ATP-EMTP

43.1 Introduction

In order to maximize power delivery in a given right of way, double-circuit transmission lines are utilized, which require taller structures than single-circuit lines which makes them more vulnerable to attract lightning; therefore, more attention should be given to analyzing their lightning performance. According to published literature based on the operating experience, lightning causes more than 50 % of the total electric power system blackouts on double-circuit transmission lines [1].

S. Wu (✉) · X. Wen · R. Li · Y. Xu
College of Electrical Engineering, Wuhan University, 430072 Wuhan, China
e-mail: whuwsh@sina.com

X. Wen
e-mail: xswen@whu.edu.cn

R. Li
e-mail: 421026731@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_43

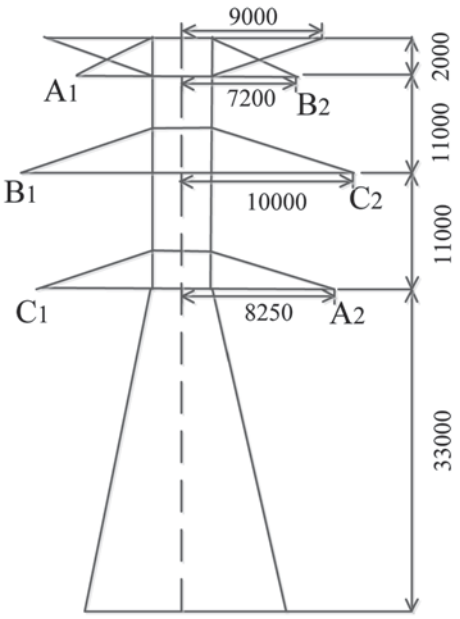


Fig. 43.1 Geometry of the tower model (dimensions are in mm)

In previous studies [2, 3], only three to five towers near the lightning stroke and the lines between them are taken into consideration to simplify the model; as a result, the length of the lines becomes much shorter. The reflected wave can seriously affect the accuracy of the results; however, it has been neglected in most researches [4]. In order to eliminate wave reflection, a matching impedance whose value is equal to the wave impedance is connected at the end of the line. ATP-EMTP and MATLAB are typically used for calculating the back-flashover lightning-withstanding level and graphics processing [5].

43.2 Calculation Models and Parameters

43.2.1 Parameters of the Tower and the Wires

The double-circuit tower model selected for the calculations is shown in Fig. 43.1 [6]. The grounding resistance ranges from 5 to 20 Ω, and the height of the tower ranges from 25 to 45 m. The soil resistivity is 1000 Ωm [7].

The transmission line uses four conductor bundles in the form of a square, with a length of 45 cm on each side. The phase arrangement is also shown in Fig. 43.1.

The insulators support the lines to have Volt-time characteristics [3], which can be expressed as in Eq. 43.1, which is fitted by MATLAB:

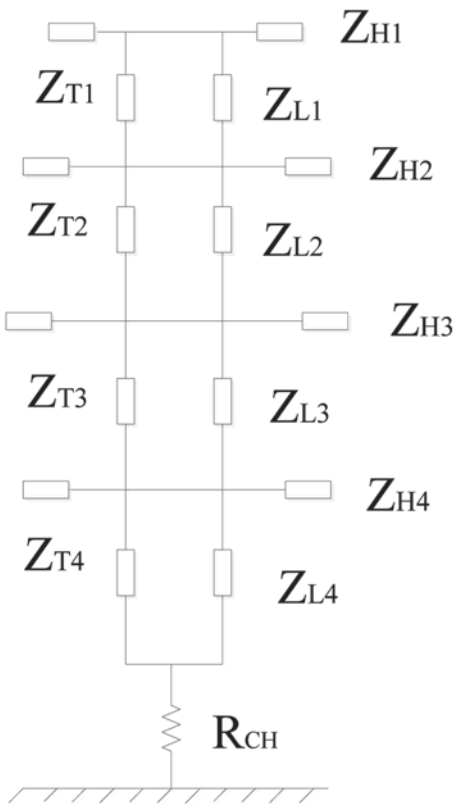


Fig. 43.2 Multi-wave impedance model of the tower

$$U(t) = 2186 + 1207e^{-t/4} + 163.8e^{-t/1.5} + 4830e^{-t/0.8}, \tag{43.1}$$

where U is in kV and t is in μs.

The schematic diagram of the tower model is shown in Fig. 43.2. Z_T, Z_H, and Z_L represent the surge impedances of main tower, cross arm, and attachments, respectively [8, 9].

43.2.2 Model Details

43.2.2.1 Lightning Current Model

There are several methods to simulate the lightning current in ATP. The surge type is adopted in this simulation, whose waveform can be described by the double exponential function [10]. The formula can be expressed as:

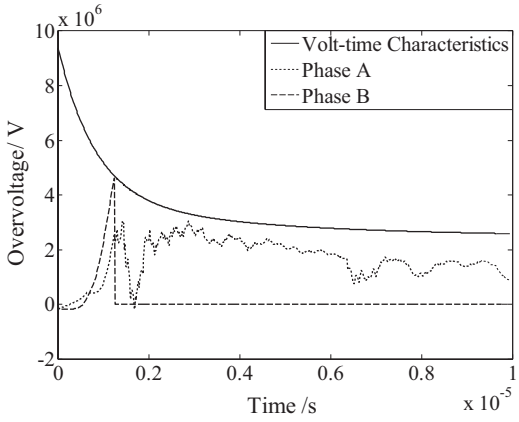


Fig. 43.3 Theory of judging the state of the insulator in ATP

$$i = I_0 (e^{-\alpha t} - e^{-\beta t}), \tag{43.2}$$

where $\alpha = -1/T_1$, $\beta = -1/T_2$. T_1, T_2 are constants of time to the half amplitude and wave front time, which decide the rise time and decay time of the lightning current.

43.2.2.2 Flashover Model of Insulator

The state of the insulator is judged by comparing its Volt-time curve with the voltage across the insulator, as shown in Fig. 43.3. If the Volt-time curve intersects the voltage curve, the insulator will flashover (phase B insulator); otherwise, it will not flashover (phase A insulator) [11].

The Volt-time curve of the insulator is mainly determined by its material characteristics and the environment. The voltage across the insulator is determined by the voltage on the tower and the overhead conductor. In Fig. 43.3, phase B is located at the top of the tower, and phase A is located in the middle; therefore, phase B suffers more severe overvoltage and flashovers than phase A.

43.3 Results and Analyses

The lightning current used for the calculations conform the international standard, whose waveform is 2.6/50 μs [10]. The basic phase arrangement is $A_1B_1C_1/B_2C_2A_2$. The height of the tower is 33 m. The grounding resistance of the tower is 10 Ω.

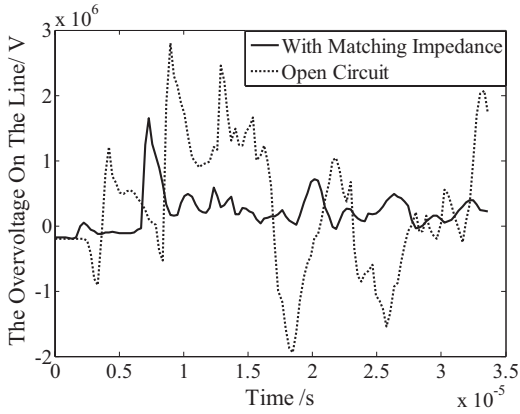


Fig. 43.4 Effect of matching impedance on line voltage

43.3.1 Analysis of the Reflected Wave

In the simulation, four towers near the lightning stroke are considered. The overvoltage on the transmission lines in the open circuit condition with the match impedance connected is shown in Fig. 43.4.

The overvoltage becomes much smoother and smaller with the matching impedance; therefore, it is important to consider the matching impedance in the simulation.

43.3.2 Overvoltage on the Tower

Table 43.1 shows the voltage on the top, the middle, and the bottom of the tower.

Table 43.1 shows that the voltage along the tower changes with time and position. The top of the tower suffers the most severe overvoltage, which indicates that the top insulator strings are most likely to flashover first.

Table 43.1 Voltage in different parts of the tower (kV)

Lightning current (kA)	Top	Middle	Bottom
80	3529	3105	2466
100	4212	3859	3074
120	4754	4011	3644

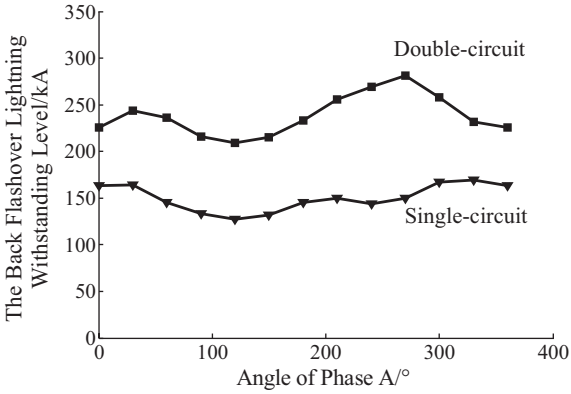


Fig. 43.5 Effect of the phase angle on lightning-withstanding level

43.3.3 Calculation of Back-Flashover Lightning-Withstanding Level

The instantaneous values of the voltage must be considered in calculating the back-flashover voltage. This voltage changes with phase angle periodically.

From Fig. 43.5, it can be concluded that the lightning-withstanding level is related to the phase angle. When the angle of phase A is about 120°, the voltage is at a minimum, and it is at a maximum when the phase angle is about 270°.

Assuming that the probability of lightning at any phase angle is equal, the lightning-withstanding level can be calculated by following statistical formula:

$$I = (\sum_{k=1}^{n} I_k) / n, \tag{43.3}$$

where n is the number of phase angle intervals in each cycle (taken as 12).

43.3.4 Factors Affecting the Back-Flashover Lightning Trip-Out Rate

43.3.4.1 Grounding Resistance of the Tower

In order to analyze the impact of grounding resistance on the back-flashover lightning trip-out rate, the grounding resistance is considered to be equivalent to a single impedance. By changing the value of the impedance from 5 to 20 Ω, the back-flashover lightning trip-out rate changes, as shown in Fig. 43.6.

It can be concluded that both the single- and the double-circuit trip-out rates increase when the resistance increases. The single-circuit trip-out rate is more sensitive when the grounding resistance is smaller than 7 Ω. The back-flashover trip-out

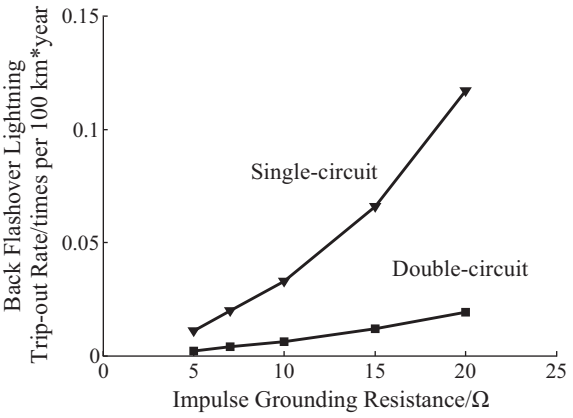


Fig. 43.6 Effect of the grounding resistance on back-flashover trip-out rate

rate is under 0.02 times/100 km*year, which is acceptable. When the grounding resistance is larger than 10 Ω, the curve becomes steep and not acceptable as a result. The back-flashover trip-out rate of double-circuit line is less than 0.019 times/100 km*year.

Therefore, it is important to maintain low values of ground resistance.

43.3.4.2 Height of the Towers

The height of the towers ranges from tens of meters to more than 100 m. By changing the height of the tower, the back-flashover lightning trip-out rate changes as shown in Fig. 43.7. The horizontal axis indicates the distance of the lowest phase conductor to the ground.

There are two reasons leading to these results: (1) The inductance increases with the height of the tower; (2) the time for the lightning current flowing through the tower correlates to the height. The taller the tower is, the more time it will take and the higher voltage may appear on the top of the tower.

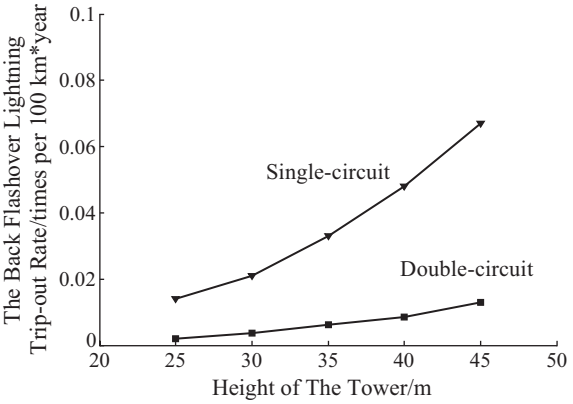


Fig. 43.7 Effect of the height of the tower on back-flashover trip-out rate

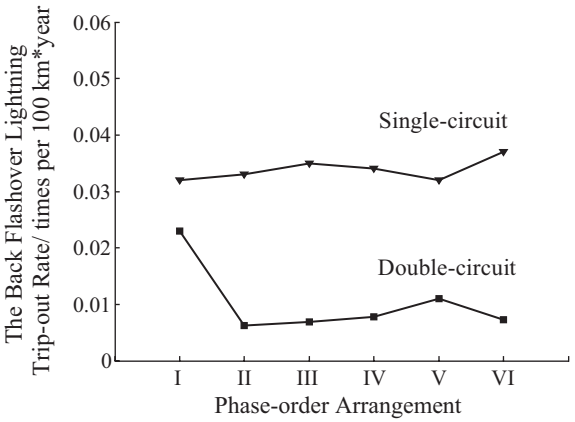


Fig. 43.8 Effect of the phase arrangement on back-flashover trip-out rate

43.3.4.3 Phase Arrangement

As to the double-circuit transmission lines, there are six possible phase arrangements: $A_1B_1C_1/A_2B_2C_2$ (I), $A_1B_1C_1/B_2C_2A_2$ (II), $A_1B_1C_1/A_2C_2B_2$ (III), $A_1B_1C_1/B_2A_2C_2$ (IV), $A_1B_1C_1/C_2A_2B_2$ (V), and $A_1B_1C_1/C_2B_2A_2$ (VI). Figure 43.1 shows the phase arrangement $A_1B_1C_1/B_2C_2A_2$, which indicates that the phase-order of the circuit on the left is ABC and the circuit on the right is BCA, from top to bottom.

Figure 43.8 gives the single- and double-circuit back-flashover rates in above arrangements. The phase arrangements have little impact on single-circuit back-flashover lightning trip-out rate. The value ranges from 0.31 to 0.37 times/100 km*year. The effect of phase arrangement on double-circuit back-flashover trip-out rate is more obvious. When the arrangement is $A_1B_1C_1/A_2B_2C_2$, the double-circuit back-flashover trip-out rate is 0.27 times/100 km*year, which is much larger than the other arrangements.

When both circuits have phase A on the top where flashover is most likely to occur, the back-flashover lightning-withstanding levels of both circuits are the lowest; therefore, the arrangement of $A_1B_1C_1/A_2B_2C_2$ should be avoided to improve the lightning-withstanding performance.

43.4 Conclusion

- It is important to consider the matching impedance in the simulation.
- The instantaneous voltage should be considered in the calculations. A statistical method can be adopted to take the results in a cycle into consideration.
- The grounding resistance and the tower height show a positive correlation with back-flashover lightning trip-out rate. The single-circuit trip-out rate increases higher and increases more rapidly than the double-circuit trip-out rate.
- The phase arrangement of A1B1C1/A2B2C2 should be avoided. In this situation, the double-circuit trip-out rate is much larger than the other arrangements.

References

1. IEEE Working Group on Estimating Lightning Performance of Transmission Lines. A simplified method for estimating lightning performance of transmission lines. IEEE Trans Power Deliv. 1985;4(4):919–32.
2. Yamada T, Mochizuki A, Sawada J, et al. Experimental evaluation of an UHA tower model for lightning surge analysis. IEEE Trans PD. 1995;10(1):393–402.
3. Qiu D, Wang A. Research on lightning protection of 500 kV double circuit transmission lines. J North China Inst Water conserv Hydroelectr Power. 2009;30(2):57–9. (In Chinese).
4. Sun Y, Sima W, Yang Q, et al. Lightning performance of quadruple-circuit transmission line with dual voltage 1000 kV/500 kV on the same tower. High Voltage Eng. 2011;37(9):2102–11. (In Chinese).
5. Prikler L, Hoidalen HK. ATPDraw version 3.5 for windows 9x/NT/2000/XP users' manual. Trondheim: INTEF/EFI; 2002. p. 28–33.
6. Wu P. Study on lightning protection performance for 500 kV double circuit transmission line. GuiZhou: GuiZhou University; 2005. (In Chinese).
7. Xie G. Hyper-voltage of power systems. Beijing: Water Power Press; 1988. p. 178–212. (In Chinese).
8. Hara T, Yamamoto O. Modeling of a transmission tower for lightning-surgeanalysis. IEE PGTD. 1996;143(3):283–9.
9. Xu Y, An Y. Study on prevention of UHV transmission. J Res Anal. 2011;39(3):37–40. (In Chinese).
10. Guan G. High voltage engineering fundamentals. Beijing: Electric Power Press; 2003. p. 113–5.
11. Zhang Y, Zeng R, Zhang B. Study on phase-order arrangement of 750 kv double circuit transmission line in same tower. Shanxi Electr Power. 2008;36(6):1–5. (In Chinese).

Chapter 44
Fault Diagnosis for Power Transformer Mainly Based on Power-On State Variables

Jinjiao Kong, Chunyu Yan, Feng Wang, Zhongyuan Zhang and Honggang Chen

Abstract Nowadays, the synthetic diagnostic methods using both power-on and power-off state variables are widely studied and applied. The usage of multiple state variables can increase the accuracy of diagnosis, but it usually brings many problems into practical application at the same time. The main problem is that the more the power-off state variables, the longer the time without power, and the worse the reliability of transformers. Therefore, this chapter presents a diagnostic method mainly based on power-on state variables. It consists of primary diagnosis and precise diagnosis. Primary diagnosis uses only power-on state variables to acquire a part of the fault information. Then, in order to get the exact fault types and more position information, based on the results of primary diagnosis we select and use some useful and necessary power-off state variables in the precise diagnosis stage. And during this process, the rough set theory is used to predigest knowledge and reduce the complexity of diagnosis. Compared with other diagnosis methods, this method shortens the outage time and improves the efficiency of diagnosis. In the end, the diagnosis is proved to be practical and effective by a fault case.

Keywords Power transformer · Fault diagnosis · Primary diagnosis · Precise diagnosis · Rough set theory

44.1 Introduction

In recent years, the research about all kinds of comprehensive diagnostic methods using dissolved gas analysis (DGA) data, oil test data, electrical test data, and inspection data is being carried out. The introduction of the electrical test data greatly increases the accuracy of diagnosis, but also brings some problems at the same time.

J. Kong (✉) · C. Yan · F. Wang
China Electric Power Research Institute, 100192 Beijing, China
e-mail: kongjj2009@163.com

Z. Zhang · H. Chen
Gansu Electric Power Corporation, 750050 Lanzhou, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_44

The main problem is that some of the data can only be acquired when transformers are powered off, so it will extend transformers' outage time and reduce the operational reliability. Therefore, this chapter proposes a diagnosis strategy which can reduce the shutdown time appropriately.

At present, many artificial intelligence methods have been applied to the synthetic fault diagnosis for transformers, such as artificial neural network, expert system, rough set theory (RST), and support vector machine (SVM), etc [1]. The diagnostic methods based onthe RST can simplify knowledge and discover the hidden relationship between data. So it contributes to synthetically applying all kinds of state information to fault diagnosis. Therefore, this chapter uses the RST in the diagnosis process.

44.2 Rough Set Theory

The RST is a new mathematics tool to deal with vagueness and uncertainty. It can discover implicit knowledge and open out potentially useful rule by efficiently analysing and dealing with all kinds of imprecise, incomplete, and disaccord information.

In RST, the knowledge expression system S can be denoted as $S = (U, A, \{Va\}, f)$ where U is the domain, A is an attribute set, and Va is a set of attribute values. $f: U \to Va$ is an information function. The "attribute-value" forms a two-dimension table, and it is an information table. If $A = C \cup D$, C is condition attribute, and D is result attribute, then a decision table is formed. The decision table can be considered as a group of defined equivalence relation that is knowledge based.

In RST, if the condition part of the rules is expressed by the condition attributes of knowledge system and the conclusion part is expressed by the decision attributes, then each object can be easily expressed to a production rule. But not all condition attributes in the original decision table are necessary. Some of them are superfluous, and removing these attributes does not affect the original expression's effect. Therefore, when describing the knowledge with rough set, reducing the decision table by removing redundant condition attributes can simplify the knowledge and transform a complicated decision table into the simplest form, which excludes redundant attributes and can also ensure correct classification of the decision table [2, 3]. The reduced decision table is an incomplete decision table. It only contains those necessary condition attribute values in decision-making, but it has all knowledge in the original knowledge system.

And the so-called solution of minimum condition attributes is just as same as solving relatively positive region in math. This chapter uses MATLAB programming to get the relative positive region [4].

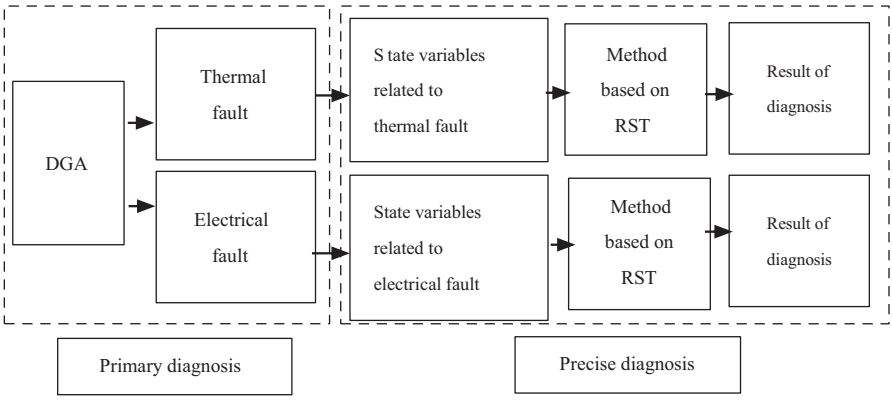


Fig. 44.1 The structure diagram of the diagnosis model

44.3 Structure and Principle of Diagnosis Model

This chapter establishes a diagnosis model which uses a variety of state variables and it can get exact fault information step by step. The structure chart of the integrated diagnosis model is shown in Fig. 44.1. In the primary diagnosis stage, the faults are divided into two categories by DGA: overheating and discharge. Then, according to the result of primary diagnosis, we choose those relevant, useful, and necessary test data for the careful diagnosis by using the RST to simplify knowledge and extract rules.

This model only uses DGA data in the first stage. So we can get the early results when transformer is powered on. Then in the second stage, only the most important and necessary power-off data need to be used for the further diagnosis. Therefore, on the one hand, because of the usage of multiple state variables, this model is more accurate than the methods in which only DGA data are used. And on the other hand, it can improve the traditional diagnostic method by using less power-off data and then shorten transformers' power outage time.

44.4 Establishing the Diagnosis Model

44.4.1 Primary Diagnosis

In primary diagnosis, the input data are the amount of dissolved gas in transformer oil. Starting conditions of the whole diagnostic process contain the amount of H_2, C_2H_2, $\sum C$, and the gas production rate of total hydrocarbon. If more than one condition exceeds the standard, we will start the diagnosis process.

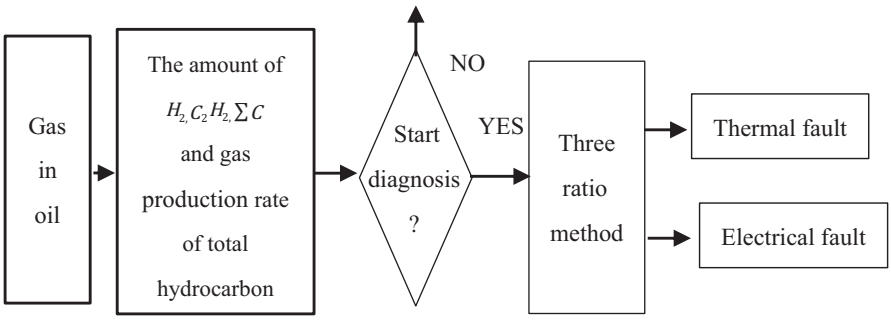


Fig. 44.2 Flow chart of primary diagnosis

When the diagnosis starts, according to the three-ratio method, a judging about the fault type is made by dividing it into thermal fault and electrical fault (Fig. 44.2).

44.4.2 Precise Diagnosis

44.4.2.1 Determining the Condition Attributes Set and Decision Attributes Set of Thermal Fault

By summarizing and classifying transformers' common symptoms and fault types, we select information related to thermal fault to create the condition attributes set *C1* and decision attributes set *D1* (Table 44.1).

Table 44.1 Condition attributes set *C1* and decision attributes set *D1* of thermal fault [5]

C1		D2	
C11	Core earth current	D11	No faults
C12	Unbalance rate of winding direct current (*DC*) resistance	D12	Core earthing or short circuit
C13	Water content in transformer oil	D13	Insulation aging
C14	Action of the gas relay protection	D14	Overheating caused by leakage flux or discharge on barrier
C15	CO_2/CO	D15	Turn insulation damage and inter-turn short circuit
C16	Deviation of voltage ratio	D16	Insulation moistening
C17	Electric capacity between windings or windings and ground	D17	OLTC(on-load tap changer) or leads faults
C18	Insulation dielectric loss	D18	Winding distortion and inter-turn short circuit
C19	No-load loss	D19	Bare metal overheating
C110	The absorption of winding	D110	Overheating because the oil flow is blocked

Table 44.2 Condition attributes set *C2* and decision attributes set *D2* of electrical fault [5]

C2		D2	
C21	Partial discharge (PD)	D21	No faults
C22	CO_2/CO	D22	Discharge on barrier
C23	Action of the gas relay protection	D23	Suspended discharge
C24	Water content in transformer oil	D24	Discharge in magnetic circuit

44.4.2.2 Determining the Condition Attributes Set and Decision Attributes Set of Electrical Fault

We create the condition attributes set *C2* and decision attributes set *D2* of electrical fault as 4.2.1 (Table 44.2).

44.4.2.3 Establishing Decision Table for Thermal Fault and Electrical Fault

The values of condition attributes are usually continuous. And according to the RST, we have to discretize them. Before discretization, in order to make all kinds of data comparable, they need to be normalized or standardized to 0–1. And it indicates the state is better if the normalized data are closer to 0. On the contrary, if it is close to 1, then we can infer that this state variable probably exceeds normal level. Then according to commonly used discrete method based on rough set, when $Cij > 0.5$, we write it to 2 in the decision table, and $Cij < 0.5$ to 1, $Cij = 0$ to 0. In this way, each condition attribute can be discretized. So the condition attributes' code can be 2, 1, and 0. They can be construed so that the probability of failure caused by this symptom is large, small, or almost none, respectively [6].

By collecting and analysing the transformer fault cases and related test data, we create decision tables for thermal fault and electrical fault as shown in Table 44.3 and 44.4.

Table 44.3 Decision table for thermal fault

C11	C12	C13	C14	C15	C16	C17	C18	C19	C110	
0	0	0	0	0	0	0	0	0	0	D11
2	0	1	1	0	0	0	0	2	1	D12
0	0	1	0	2	0	0	1	0	1	D13
0	0	0	1	0	0	0	0	1	1	D14
0	1	0	1	1	0	1	0	0	1	D15
0	1	1	0	0	0	0	0	0	0	D16
0	2	1	1	0	1	1	1	0	0	D17
1	0	0	1	0	0	1	0	0	2	D18
0	0	1	1	1	0	0	0	1	0	D19
0	0	1	0	0	0	0	1	1	0	D110

Table 44.4 Decision table for electrical fault

C21	C22	C23	C24	
0	0	0	0	D21
2	0	1	0	D22
2	2	1	1	D23
1	0	1	1	D24

44.4.2.4 Reducing Decision Tables

By using the programme written in MATLAB for solving the relative positive region, Table 44.3 and 44.4 are reduced and the minimalist condition attributes sets are gotten. But considering the difficulty and accuracy of the test data, we choose {C11, C12, C13, C15, C19} as the final decision of C1. For the same reasons, finally, we choose {C21, C22, C23} as the minimal condition attributes set of *C2* [7].

44.4.2.5 Diagnosing by Using the Simplest Decision Tables

After getting the simplest decision tables, we can apply them to fault diagnosis by using matching methods. And this chapter uses the simplest matching method—Euclidean distance method. In this method, the Euclidean distance d_{0i} is the distance between the target reduction and source reduction. The smaller d_{0i} is, the better they can match.

$$d_{0i} = \sqrt{\sum\nolimits_{k=1}^{n} (a_{0k} - a_{ik})^2}, \tag{44.1}$$

where a_{ik} is the number in *i*th row and *k*th column in the decision table.

After comparing every d_{0i}, the corresponding fault type D1*i* or D2*i* of the smallest d_{0i} can be considered as the actual fault type.

44.5 An Example of Fault Diagnosis

On 6 July 2011, it was found that the total hydrocarbon of #2 main transformer was beyond the normal value in a 35 kV substation. The DGA data of continuous tracking are showed in Table 44.5.

According to the diagnosis model established in this chapter, firstly, from the Table 44.7, we find that the amount of H_2, C_2H_2, and the total hydrocarbon all exceed the normal range. Absolute producing rate of them are 122.75, 3.31, and 659.87 μL/d, respectively. And relative producing rate of the total hydrocarbon is 136.94% per month. They are all beyond the attention values. Therefore, the diagnosis process begins. In the primary diagnosis stage, according to the three-ratio

Table 44.5 Dissolved gas analysis (DGA) data of #2 main transformer

Date of test	Test reason	Data(μL/L)							
		H_2	CO	CO_2	CH_4	C_2H_6	C_2H_4	C_2H_2	Total hydrocarbon
2011.07.06	Normal	66.43	88.26	1645.96	158.77	52.19	303.99	0	514.95
2011.08.03	Tracking	70.29	100.78	1506.38	203.31	83.38	468.18	0	754.87
2011.09.26	Tracking	88.72	86.24	1428.45	252.55	112.19	625.97	1.46	992.17
2011.10.24	Tracking	106.25	91.35	1480.44	281.98	129.30	687.72	2.56	1081.56
2011.11.22	Tracking	142.50	123.66	1580.44	296.26	142.08	727.45	3.08	1168.87
2011.12.03	Tracking	676.74	176.11	1665.02	969.55	570.57	2483.26	17.48	4040.86

Table 44.6 The values of condition attributes in the example

Simplest condition attribute	C11	C12	C13	C15	C19
Test data	0.1A	2.4%	22 g/L	9.45	No change
Code	0	2	1	1	0

Table 44.7 The Euclidean distance between the example code and the target code

d01	d02	d03	d04	d05	d06	d07	d08	d09	d010
2.45	3.16	2.45	2.65	1.73	1.41	1	3.32	2	2.24

method, the code is 022, so it belongs to the category of thermal failure [8]. Then precise diagnosis begins. In this stage, the transformer needs to be stopped and tested. The measured unbalanced rate of three-phase DC resistance is 2.4%, and the earthing current of transformer core is 0.1A. From the oil test, we get the water content in oil is 22 g/L. And compared with the value tested when leaving the factory, the no-load loss has no obvious change. According to Table 44.5, CO_2/CO ratio is 9.45. The values of condition attributes are shown in Table 44.6 below.

Then, according to Table 44.3, by using the equation (44.1), we calculate the Euclidean distances between the example code and the target code.

According to the result, we can easily find that d_{07} is minimal. It means that the example's fault type is most similar to *D17*, so the fault can be diagnosed as overheating caused by the defects of tap-changer and leads. Finally, by lifting core inspection, we found that the transformer tap-changer static contact screw was loose and the interface of static contact was burning and black. So it proves that the diagnosis result is correct and useful.

44.6 Conclusion

In consideration of the shortage of the existing diagnosis, a two-step method mainly based on power-on state variables is given in this chapter. Firstly, the primary diagnosis is made based on the limited power-on state variables to get partial fault information. Then precise diagnosis is carried out by using some necessary and useful power-off state variables which are chosen according to the results of primary diagnosis. In the precise diagnosis, in order to reduce the complexity of diagnosis, the RST is used to simplify knowledge and extract rules. The diagnostic method designed in this chapter can shorten blackout time of transformer to some extent and improve the efficiency of the diagnosis and the reliability of transformers. But it still cannot complete the diagnosis with transformers in service all the time, so better diagnosis method is yet to be developed.

References

1. Saha TK. Review of modern diagnostic techniques for assessing insulation condition in aged transformers. IEEE Trans Dielectr Electr Insul. 2003;10(5):903–17.
2. Bai J. Study on the application of rough set theory in the transformer fault. Tianjin: Tianjin University; 2009 (in Chinese).
3. Wang Y, Lu F, Li H. synthetic fault diagnosis method of power transformer based on rough set theory and bayesian network. Proc CSEE. 2006;26(8):137–41 (in Chinese).
4. Cheng Y, Chen Z. The realization about algorithm of relative positive region. Comput Appl Softw. 2005;22(3):30–1,82 (in Chinese).
5. DL/T 596–2005. Preventive test code for electric power equipment (in Chinese).
6. Wang N, Lu F, Liu Y, et al. Synthetic fault diagnosis of oil-immersed power transformer based on rough set theory and fuzzy Petri nets. Proc CSEE. 2003;23(12):128–32 (in Chinese).
7. Huang W, Zhao X, Wang W et al. Extraction method of decision rules for fault diagnosis based on rough set theory. Proc CSEE. 2003;23(11):150–4 (in Chinese).
8. GB7252–2001. Guide for the analysis and diagnosis of gases dissolved in transformer oil (in Chinese).

Chapter 45
The Ice Edge Detection of Transmission Line Based On Structured Forest Algorithm

Zhenyu Wang, Jianfeng Jia and Jing Teng

Abstract Focusing on the existing problems of accuracy and latency in the current ice transmission line edge image detection, this chapter proposes a machine learning algorithm based on structured forest edge detection of ice transmission line. The ice transmission line image information's model is automatically trained by the structured forest algorithm and the model is automatically applied for edge detection of the ice transmission line. Experimental results show that the method of ice transmission line edge detection is not only feasible and effective, but also can accurately detect ice line edge profile. What is more, it can meet the requirement of real time detection compared with other algorithms. Satisfactory results are obtained in the test.

Keywords Structured forest · Ice transmission line · Edge detection

45.1 Introduction

With the influences of micro climate, terrain, and weather condition, more and more countries suffer more from ice disaster which occurs frequently. And by the increasing of power load, power grid reliability becomes more and more important [1]. However, the transmission line ice cover largely poses a threat on the reliability of the power grid which can cause serious harm accidents to the safe operation of power system, such as the flashover of insulator string, rod, etc. These give rise to huge economic losses [2]. Therefore, it is of great importance to detect ice transmission line edge accurately and in real time.

Edge detection is the most basic area in the image processing and computer vision technology. Early classical algorithms include edge operators, curved surface fitting, matching method, threshold method, and other methods [3]. In recent years, with the development of mathematical theory and artificial intelligence, many new methods

Z. Wang (✉) · J. Jia · J. Teng
School of Control and Computer Engineering, North China Electric Power University, 102206 Beijing, China
e-mail: zywang@ncepu.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_45

emerge, such as wavelet transform and wavelet packet method, mathematical morphology, fuzzy theory, and neural network method [4]. These traditional methods mostly adopt the traditional edge detection operator and the method of threshold segmentation using a large number of color contour edge information to edge detection. Its color grayscale does not match with many prominent edges that cause detection precision. In the traditional edge detection methods, we should normalize the inputs before edge detection. But input normalization process is time consuming. So it cannot satisfy the requirement of real time [5]. Therefore, this chapter proposes a new method which adopts the structured forest algorithm for detecting the ice edge of the transmission line. The structured learning method is used to calibrate the output variables, and then we can obtain the structured information by using the method of random forest algorithm, which is used for ice transmission line model training and edge detection. The proposed method can meet the accuracy and real time requirements that the traditional methods cannot meet, so it has a great application value [6].

45.2 Line Edge Detection Based on the Structured Forest Algorithm

According to the characteristic of ice transmission line edge blurred, the structured forest algorithm is adopted to detect the edge of the image in this chapter. The structured forest algorithm is mainly based on the random forest algorithm combined with the structured learning theory. The edge detection precision and real-time performance have a very big enhancement according to the experiments, especially in the processing speed of the algorithm.

45.2.1 Random Forest Algorithm

Breimann put forward random forest algorithm which is a method composed of decision trees [7]. Each tree based on a random variable is an independent set value and independent sampling. What is more, they obey a uniform distribution. Each tree represents a pixel edge in the image and each tree's node is a binary segmentation function [8]. In this chapter, we select the images that contain multiple channels in image processing. For example, First Red Green Blue (FRGB) or red green blue depth (RGBD) images, etc. As it is very complex and has a large amount of calculation, so we use:

$$h(x,\Theta) = [x(k_1) - x(k_2) \le \tau], \tag{45.1}$$

where, x is the information variable of input images; $\Theta = (k_1, k_2, \tau)$, and Θ is the random variable of each tree, in which $x(k_j)$ represents the pixel of each decision tree; τ represents the pixel's threshold, while $h(x,\Theta)$ represents the pixel value. The pixel point is considered to the edge of the image only if the pixel difference

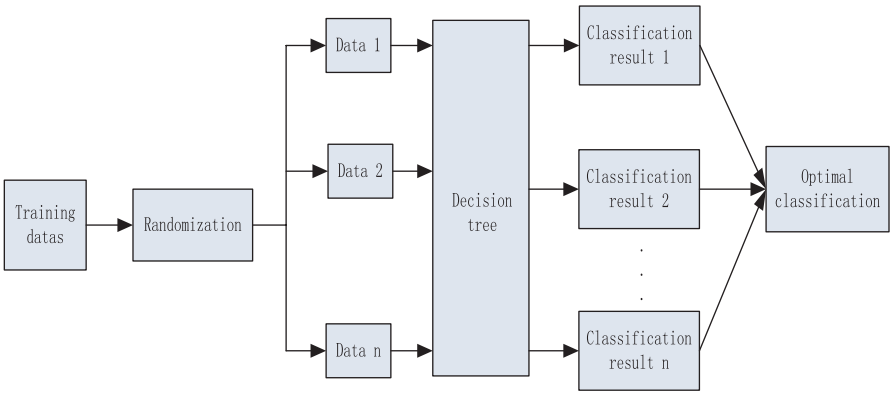


Fig. 45.1 Random forest algorithm

between the two points is within the scope of the threshold. The random forest algorithm is a system composed by growing decision trees[9], as shown in Fig. 45.1:

45.2.2 Edge Detection of Structured Forest Algorithm

Structured forest algorithm is formed by the junction of random forest algorithm and structured learning in which structured learning is mainly used for the structured definition of all kinds algorithms of inputs and outputs in a linear discrete space. The purpose of putting two algorithms together by the mean of structured learning is to achieve the processing data rapidly, which can accelerate the speed of the algorithm to perform. In this chapter, due to the reason of the speed of the algorithm performing, only the output variables are defined in structure, which not only ensure the accuracy of the detection, but also improve the speed of the edge detection.

The structured forest algorithm's major concept is that by adopting the method of structured learning to put a pixel and its distribution points together to form a regional Z in the random forest value of Y, which can guarantee the real time of ice transmission line edge detection, as shown in the formula (45.2):

$$\prod : Y \to Z. \tag{45.2}$$

According to the need of the actual algorithm, Y is a structured output value of containing 16×16 pixels in this chapter and Z is a region in the random forest. Through the formula (45.2) processing, it will transform each pixel's information in the area of 16×16 into Z which is a long binary variable containing each pixel's binary coded information. As transforming the information of color image to a binary information through the transformation (45.2), it is convenient for computer processing and it improves the running speed. But Z is a multi-dimensional

vector, so we need to reduce the dimension of Z to guarantee the speed of the program running under the premise of transmission line edge information preserved. Methods to reduce the dimension of Z at present are mainly two kinds: One kind is the K—means method, while another is to use the method of principal component analysis (PCA). But, reducing the dimension needs to meet the following criteria:

$$p^* \le \frac{p(1-s^2)}{s^2}. \tag{45.3}$$

Among them: p^* is the generalization error of the structured forest; s is the strength of tree classification, and p is the average correlation coefficient between the trees. In the process of reducing the dimension of Z, we must make the pledge that the generalization error is less than the maximum value in the structured forest algorithm [10].

Therefore, the ice transmission line detection based on the structured forest is that we put the ice transmission line images as inputs to extract the structured information and texture information, and then automatically by the model of training and random forest edge detected, we can get the output information of ice line edge. At last, the final ice transmission line edge detection image is acquired by using the method of structured learning. In the process of simulation, we use the particle swarm optimization (PSO) algorithm and structured forest algorithm for image edge detection. The two kinds of algorithms are evaluated from the edge detection of the inosculation and simulation time, as shown in Figs. 45.3 and 45.4.

45.2.3 Model Training

Each of the decision trees in structured forest algorithm can be seen as a weak classifier, which is trained by using a recursive method independently. To the image training sample set Ω about ice transmission lines that reach the node, a classification criterion $h(x,\Theta)$ is calculated, as the formula (45.4):

$$h(x,\Theta) = \delta(\varphi^T(x)\psi(x) > 0). \tag{45.4}$$

Where, $x \in R^M$ is set to the training sample and $\Theta = [\varphi,\psi]$ presents the weak classifier's parameter, among which $\Theta(\bullet)$ served as the filter function. What is more, ψ is a parameter matrix or column. The Θ determines the classification hypersurface form of a weak classifier; δ is an indicator function. When the sample meeting $h(x,\Theta)$ is less than the setting threshold, it falls into the left node; on the other hand, it falls into the right node. We should do the recursive of the above process until the numbers of samples falling into the node less than the threshold or the sample purity

is greater than the threshold. At this point, the node becomes a leaf node. Otherwise, we should find the optimal coefficient Θ^* in each node to make the information gain maximization after training sample classification [11]. The ice transmission line classification and the model are acquired after training all leaf nodes according to the rules of combination together.

45.3 The Experimental Results and Analysis

In order to test the algorithm's running speed when running the project in this algorithm, we specially choose the ice transmission line in the fog weather. We make image enhancement before the edge detection. According to the actual situation, this chapter selects the histogram equalization to image enhancement [12]. The results are shown in the Fig. 45.2 below:

According to the above ice transmission line image after image enhancement, we make two kinds of algorithms run in MATLAB by using the runtime environment for 2 core@2.2Ghz CPU and 2 GB memory of the computer. The two kinds

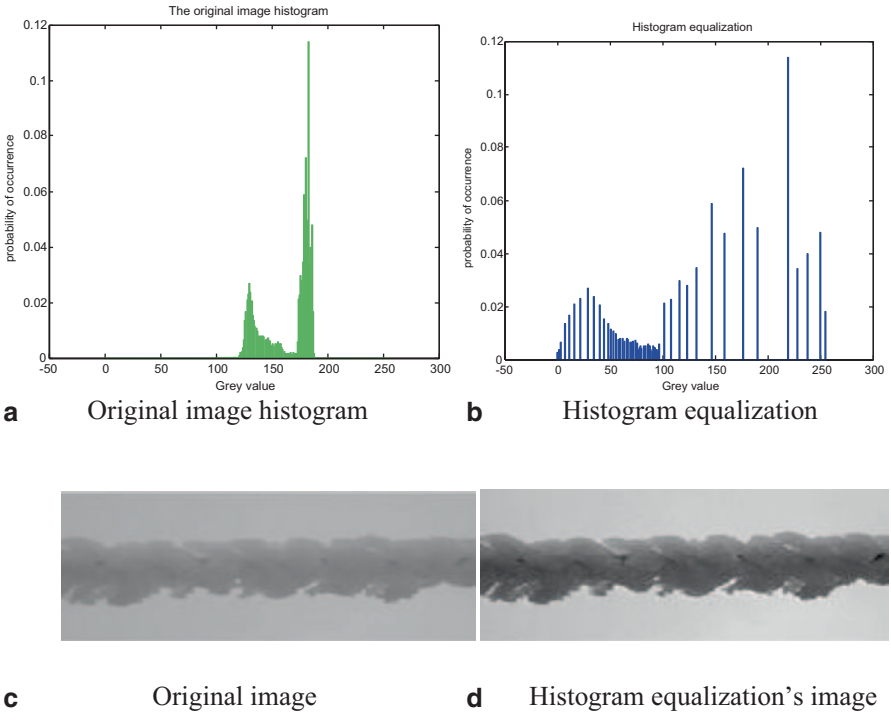


Fig. 45.2 Image enhancement comparison a Original image histogram. b Histogram equalization. c Original image. d Histogram equalization's image

of algorithms include the PSO and structured forest edge detection method. Two groups of the contrast experiments are carried and experimental results are as follows:

45.3.1 Experimental Contrast Figures

In the experimental process, we change detection rate by the structured forest algorithm and get the edge detection images which are shown in the Fig. 45.3. By the experimental results, it can be seen that the system running time is increasing in the case of the pursuit of high detection precision. So in the actual situation, we need to choose in terms of detection accuracy and program running real time. But compared with other methods, the structured forest method not only improves the accuracy of edge detection, but also it reduces the running time (Table 45.1).

45.3.2 Algorithm Experimental Results

In the above operation environment, we put two kinds of algorithms in MATLAB by operating the two groups of enhancement pictures and the results are shown in the Fig. 45.4. By edge detection results, it can be seen that the particle swarm algorithm could not fully detect the edge of the ice transmission line. What is more, the processing time is much greater than the structured forest algorithm. Meanwhile, the structured forest algorithm not only can accurately detect the edge of the ice transmission line, but also the program is running fast and convenient for measuring the diameters of the ice transmission line compared with other algorithms (Table 45.2).

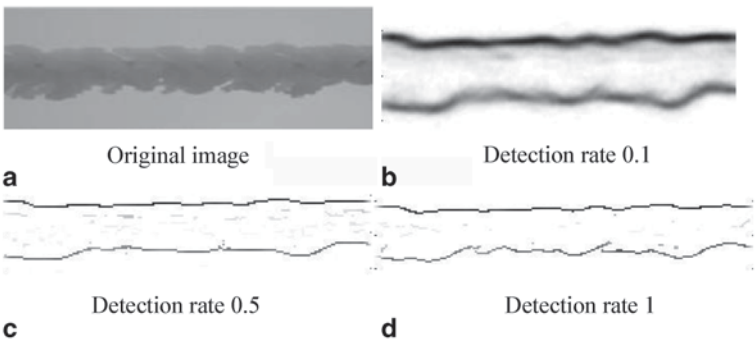


Fig. 45.3 The simulation contrast figures. **a** Original image. **b** Detection rate 0.1. **c** Detection rate 0.5. **d** Detection rate 1

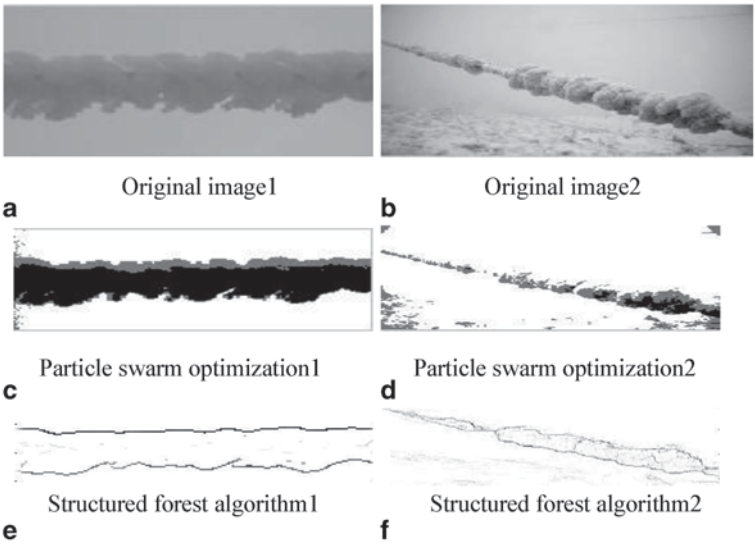


Fig. 45.4 Contrast figures. **a** Original image1. **b** Original image2. **c** Particle swarm optimization1. **d** Particle swarm optimization2. **e** Structured forest algorithm1. **f** Structured forest algorithm2

Table 45.1 Processing time1

Name		Enhance image
Size		181*74
Time/s	(b)	0.051
	(c)	0.054
	(d)	0.059

Table 45.2 Processing time2

Name		Original image1	Original image2
Size		181*74	329*220
Time/s	PSO	0.205	0.952
	Structured forest	0.054	0.337

PSO particle swarm optimization

45.4 Conclusion

Aiming at the problem of ice transmission line detection based on the images in this chapter, we put forward the structured forest algorithm for ice transmission line edge detection. As shown in the experimental experiments, the structured forest method for ice transmission line edge detection is not only of high precision, but also the program processing speed is very fast. The method can automatically adjust

the size of the detection rate according to actual circumstances. More importantly, compared with other algorithms, this algorithm's processing time is greatly reduced and the detection rate is very high in dealing with the same image. Therefore, this method is not limited to the transmission line ice edge detection. It can also be applied to the insulator and so on, which has a higher application value in the field of electrical equipment for edge detection. But the current program for transmission line inner edge detection's effect in the process of experiment can also continue to improve in the next step to better meet the practical requirements.

Acknowledgements This work was supported in part by the Fundamental Research Funds for the Central Universities Grant.

References

1. Yi HU, Yi H. Analysis and countermeasures discussion for large area icing accident on power grid. High Volt Eng. 2008;34(2):001–2.
2. XU S, ZHAO J. Review of ice storm cases impacted seriously on power systems and de-icing technology. South Power Syst Technol. 2008; 2(2):002–3.
3. Davis LS. A survey of edge detection techniques. Computer Graph Image Process. 1975;4(3):248–70.
4. Pal NR, Pal SK. A review on image segmentation techniques. Pattern Recognit. 1993; 26(9):1277–94.
5. Papari G, Petkov N. Edge and line oriented contour detection: State of the art. Image Vision Comput. 2011;29(2):79–103.
6. Dollár P, Zitnick CL. Structured forests for fast edge detection.Computer Vision (ICCV), 2013 IEEE International Conference on. IEEE. 2013. p. 1841–48.
7. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
8. KANG Y, Nagahashi H, Sugimoto H, et al. Image categorization using scene-context scale based on random forests. IEICE Trans Inf Syst. 2011;E94-D(9):1809–16.
9. Yang BS, Di X, Han T. Random forests classifier for machine fault diagnosis. J Mech Sci Technol. 2008;22(9):1716–25.
10. Taskar B, Chatalbashev V, Koller D, et al. Learning structured prediction models: A large margin approach. Proceedings of the 22nd international conference on Machine learning (ICML), 2005 IEEE International Conference on. IEEE. 2005. p. 896–903.
11. Bowyer K, Kranenburg C, Dougherty S. Edge detector evaluation using empirical ROC curves. IEEE computer society conference on computer vision and pattern recognition, 1999. IEEE. 1999. 1:1–3.
12. Tan RT. Visibility in bad weather from a single image. IEEE conference on computer vision and pattern recognition, 2008. CVPR 2008. IEEE. 2008. p. 1–8.

Chapter 46
Lifetime Assessment and Optimized Maintenance System of Transformers Based on the HST Model

Lefeng Cheng, Bin Zhou, Dehua Cai, Weiyu Chen, Liguo Wang and Tao Yu

Abstract Currently, the maintenance mode of the power transformers is regular maintenance, which has been widely used. It is key to correctly evaluate the reliability level of the transformer in operation condition because it is directly related to the success of maintenance; on this basis, the lifetime evaluation and optimized maintenance model based on hot spot temperature (HST) in terms of the power transformer is developed in this chapter. In this model, the reaction speed theory of Arrhenius and Weibull distribution is considered as the fundamental theories, which are used to describe the aging process of the transformer and get the transformer failure rate λ; and then the equations with exponential form are adopted to compute HST, namely, to calculate the top oil temperature relative to the environmental temperature $\Delta\Theta_{TO}()$ the increment of top oil temperature $\Theta_{Ae}()$ and the hysteretic temperature $\Theta_{Ae}()$ so as to get HST (equals to the sum of the aforementioned three variables). Meanwhile, based on the model, the software analysis system based on Java language and Mysql database has also been developed, in which the transformer maintenance process was optimized and the statistical analysis was made for the hot spot area. Finally, the diagnosis results were obtained correspondingly, indicating that the designed model and system can effectively reduce the maintenance frequency, improve the utilization coefficient of the transformer, and thus improve the equipment's reliability.

L. Cheng (✉) · B. Zhou · W. Chen · T. Yu
Electric Power College, South China University of Technology, 510640 Guangzhou, China
e-mail: chenglf_scut@163.com

B. Zhou
e-mail: zhoubinscut@163.com

D. Cai · L. Wang
JiangMen Power Bureau, 529030 Jiangmen, China
e-mail: tsdd7219@21cn.com

W. Chen
e-mail: 4363095802@qq.com

L. Wang
e-mail: 13923068819@163.com

T. Yu
e-mail: taoyu1@scut.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_46

Keywords HST model · Transformer · Lifetime assessment · Maintenance optimization

46.1 Introduction

As the power transformer is an important equipment in the power system, its reliability is directly related to the security and stability of power system. Compared with the regular maintenance which has been widely used in current power system maintenance, the state maintenance method can effectively reduce the maintenance cost, shorten the outage time and improve the utilization rate of power equipment; thus, it has become the main development direction of transformer maintenance. As the correct evaluation of reliability level for the transformer in operation is the key to the success of the state overhaul, the reliability level of the transformer should be grasped in a timely and accurate manner because it is significant for guiding the transformer state maintenance, improving the operation reliability of power transformers, and thus ensuring the safe and reliable operation of the power grid [1].

For a long time, the power system primarily adopted the time-based maintenance (TBM) to judge the health level and good condition of the operation of transformers. In China, as stipulated in the national electric power industry standard "Preventive Test Rules of Electric Power Equipment" (DL/T596-1996) and "Transformer Maintenance Guidelines" (DL/T 573-95), the transformer with 110 kV and above 110 kV should be put into a large-scale maintenance when it has been operated for 5 years; upon this maintenance, the transformer should be put into a larger scale of maintenance every 10 years while a small-scale maintenance should be implemented annually. This TBM mode, which ignores the state of the transformer, has always caused "small disease for big scale of treatment or no disease for treatment" of transformers. When the equipment is not found with any fault or in normal operation, then the maintenance or replacement is still implemented, which has caused a great waste of manpower, material, and financial resources. Due to the problems of the test methods or the equipment, partial faults are not found in time; or heavy repair task or tight time results in the transformers not being timely repaired, which may cause both faults and loss in the transformers.

Some authors presented a repairable failure rate model based on condition monitoring data [2]. The maintenance strategy combining the equipment condition monitoring and the system risk evaluation is proposed by using the two models. With such a maintenance strategy, both the equipment condition and the system risk can be considered in a unified maintenance process.

With the full life-cycle cost theory used as a guide with the environmental costs and social responsibility costs taken into consideration, a more comprehensive full life-cycle cost model of the transformer was established [3]. This model can not only get economic life expectancy of the transformers but also get the economic life interval under different confidence levels and interval whose chance measurement is the greatest under the same economic life span.

At present, the research on the reliability of power transformers can be divided into the two aspects of qualitative analysis and quantitative evaluation [4]. Concerning the aspect of qualitative analysis, the China Power Grid Reliability Management Center, according to the statistical data collected over the years, analyzed the factors which may affect the reliability index of transformer operation and reliability of transformer, and also analyzed the unplanned outage cases of transformers according to the components, and finally pointed out that the transformer faults mainly concentrated in the transformer windings, casing, insulating medium, cooling device, core, tapping switch, adjusting device, oil pillow, etc. Concerning the aspect of quantitative analysis, domestic and foreign scholars have also made some research achievements. The current deeper study is to adopt the fault tree analysis mode [5] in order to analyze the reliability of power transformers according to the position and function of each part of the transformer and the fault mode of the transformer; then to indicate that the transformer should be divided into winding, casing, cooling system, core, tapping switch, insulating medium, etc.; and further to build the transformer fault tree model to realize the reliability and assessment analysis of the transformer [6–8]. Besides, there are also some other studies which divide the utilization condition of transformers into 11 kinds of state, such as normal operation, winding fault, and preventive test[9], focus on the statistical data of reliability index for power transformers, combine the Markov process theory [10], and then adopt the frequency and duration method [11] to establish the reliability evaluation model of transformer. To study the reliability of the equipment, efforts should be made not only to estimate some reliability indices such as the probability and frequency of failure but also to find out the more significant influence factors on the reliability and recognize the weak links so as to determine the measures for improving the reliability of equipment. Here, two aspects of appropriate measures of design manufacture and operational maintenance are involved. The reliability of the tracking method [12] is to use the tracking ideas to determine the distribution relationship between each element and the system reliability indexes in a scientific and reasonable manner, and then to determine the contribution of elements to the system reliability indexes. The existing literatures [12–15] on the reliability tracking theory have been used to study the reliability of electric power systems and put forward the proportion allocation criterion of reliability tracking, thus recognizing the element which most importantly contributes to the system unreliability, in which the transformer reliability tracking technology is to determine the contribution of each component of the transformer to the (possible) fault events, thus determining the key components which affect the reliability of transformers.

As to the regular maintenance of transformers, for voltage levels of transformers of 110 kV and above, it would be better to repair them on a large scale every 5 years and guarantee the implementation of the same kind of maintenance every 10 years, and a small-scale maintenance annually. The lifetime evaluation of transformers and optimization maintenance of the smart system based on hot spot temperature (HST) have been designed to obtain the failure rate of power transformers, whose failure rate is very close to that of the regular maintenance mode; thus, the designed system can play a significant role in the regular maintenance of transformers every

time. Certainly, the utilization of this developed system also does not need to be implemented according to the plan of regular maintenance. It can perform detecting analysis of the HST of transformers at any time and find the overheat and overload conditions existing in the power transformers in time as well as the isolation-aging problems and nonscheduled maintenance. At the beginning of maintenance of this newly developed system, the regular maintenance plan can be cancelled, as determined by people's expectations; furthermore, with the maturity and development of the designed system, the regular maintenance mode can be cancelled.

46.2 Structure of Transformers and Analysis of Fault Reasons

The power transformer is composed of insulating materials, conductive and magnetic materials, and structure materials [16–18], with the insulating materials consisting mostly of organic material, such as mineral oil, insulating paper, and various organic synthetic materials. Compared with the metal materials which may be affected by various factors, such as electricity, heat, and mechanical and environmental factors, when the insulating materials are in operation, they easily deteriorate gradually, significantly decreasing the electrical and mechanical strength, and further cause equipment failure and interruption of power supply. A lot of research on material show that the main factor leading to transformer accidents is deterioration of its insulating performance [14]; therefore, the insulation performance of the equipment structure is often the key factor to determine the whole lifetime of the transformers.

Generally speaking, the transformer windings are A-class insulation. Under normal loading, the insulating material can be used for more than 20 years; if it is operating upon overload, the insulation aging will be accelerated. The insulation materials will turn black upon aging and will lose elasticity and become crisp. In this case, as long as the windings are slightly affected by vibration or friction, the insulation performance may be completely damaged, which thus leads to turn-to-turn short circuits or layer short. The insulation performance will also decrease obviously upon aging. A breakdown is easy in case of overvoltage. In order to prevent and slow down the aging of insulation, we must strictly control and master the transformer load and strictly control the upper temperature and temperature rise.

The International Electrotechnical Commission (IEC) considers that as the temperature is increased by 6° C, the reduction speed of effective lifetime of transformer insulation will increase by 1 multiple when A-class insulated transformers are operating in a temperature range of 80 ~ 140° C. The lifetime of transformer depends on the aging degree of insulation, and the aging speed of insulation depends on the operating temperature.

46.3 Lifetime Assessment of Transformers Based on the HST Model

The HST model is established to calculate the failure rate of transformers, which mainly includes the following steps: the first step is to calculate the difference value $\Delta\Theta_{TO}$ between the top oil temperature and environmental temperature; secondly, it is to calculate the difference value $\Delta\Theta_H$ between HST and top oil temperature; and finally, it is to calculate the values Θ_H of winding HST; in this step, the environmental temperature is Θ_{Ae}. Institute of Electrical and Electronics Engineers (IEEE) rules point out that the insulation temperature is the control parameter of insulation aging or insulation deterioration; so the insulation lifetime L should be calculated based on the calculation of insulation temperature Θ_H, and the difference between HST and top oil temperature and solution, and the fault rate λ of the transformer can be obtained based on the lifetime of insulation. The above process is described in detail as follows.

46.3.1 Establishment of the HST Model for the Calculation of Θ_H

The transformer insulation aging failure rate is closely related to the winding HST, which mainly depends on the ambient temperature and the transformer load. The hottest spot is often located at the high-voltage winding or at the top or middle of low-voltage winding of the power transformer. If the effect of environmental temperature hysteresis is ignored, utilize the exponential form of equations to calculate HST.The calculation diagram is shown in Fig. 46.1.

The specific steps are shown as follows:

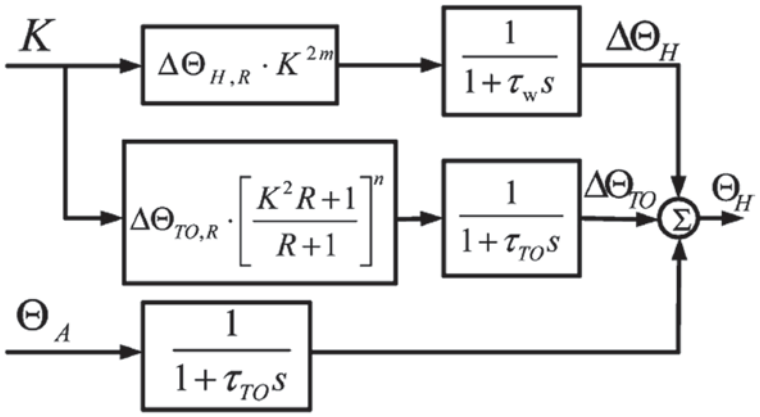


Fig. 46.1 Calculation diagram of Θ_H

Step 1: Calculate the temperature rising value $\Delta\Theta_{TO}()$ of the top oil temperature relative to the environmental temperature.

$$\tau_{TO}\frac{d\Delta\Theta_{TO}}{dt} = \Delta\Theta_{TO,U} - \Delta\Theta_{TO} \tag{46.1}$$

$$\Delta\Theta_{TO,U} = \Delta\Theta_{TO,R}\bullet\left[\frac{K^2R+1}{R+1}\right]^n \tag{46.2}$$

In these formulas, $\Delta\Theta_{TO,U}()$ is the final rising temperature of top oil relative to ambient temperature, and $\Delta\Theta_{TO,R}()$ is the rising temperature of top oil under the rated load situation. $K(p.u)$ is the ratio of actual load and rated load, $R(p.u)$ is the ratio of load loss and no-load loss under the rated load condition, and $\tau_{TO}(h)$ is the time constant of the transformer oil. $\Delta\Theta_{TO,R}()$ can be obtained by the following two ways: (a) according to the IEEE standard for the actual test, or (b) the calculation value is supplied by the equipment manufacturers in the test report.

Step 2: Calculate the increment $\Delta\Theta_H$ of winding HST related to the top oil temperature of transformer.

$$\tau_w\frac{d\Delta\Theta_H}{dt} = \Delta\Theta_{H,U} - \Delta\Theta_H \tag{46.3}$$

$$\Delta\Theta_{H,U} = \Delta\Theta_{H,R}\bullet K^{2m} \tag{46.4}$$

In these formulas, $\Delta\Theta_{H,U}()$ is the final increment of the highest temperature point relative to the top oil temperature, $\Delta\Theta_{H,R}()$ is the rising temperature value of the highest temperature point related to the oil temperature at the rated load, $\tau_w(h)$ is the time constant of the temperature point position, and m and n are the empirical constants depending on the cooling mode. $\Delta\Theta_{H,R}()$ and $\Delta\Theta_{A,R}()$ can be obtained by the following three ways: (1) by the test of the built-in temperature detector of the transformer, (2) from the equipment manufacturers in the test report, or (3) supposing the average temperature of the windings is 65° C, $\Delta\Theta_{H/A,R} = 80$, and supposing the average temperature of the windings is 55° C, $\Delta\Theta_{H/A,R} = 65$. The relationship between the value of m and n and the cooling mode of the transformer is shown in Table 46.1.

Step 3: Calculate the hysteretic environmental temperature $\Theta_{Ae}()$.

$$\tau_{TO}\frac{d\Theta_{Ae}}{dt} = \Theta_A - \Theta_{Ae} \tag{46.5}$$

In this formula, $\Theta_A()$ is the instantaneous temperature.

$$\Theta_H = \Theta_{Ae} + \Delta\Theta_{TO} + \Delta\Theta_H \tag{46.6}$$

Table 46.1 Corresponding values of *m* and *n* under different cooling modes

Type of cooling mode	*m*	*n*
OA	0.8	0.8
FA	0.8	0.8
Nondirected FOA or FOW	0.8	0.9
Directed FOA or FOW	1.0	1.0

The *O* in *OA* or *FOA* or *FOW* means the insulating liquid mineral oil or ignition temperature more than 300 °C; the *A* means air; the *W* means water; the *F* in *FOA* or *FOW* means the oil flow in the cooling equipment is forced circulation, and the oil flowing through the winding internal is the thermal convection loop

46.3.2 Establishment of the Aging Failure Model Based on HST for the Calculation of the Failure Rate

The aging fault of transformers has something to do with the lifespan of materials. Under normal load situation, the insulation strength of insulated conductors deteriorates slowly for a long period; therefore, the mechanical strength loss of insulated materials is the main reason for the equipment aging fault. It is an irreversible process as the time accumulates. Wilbur distribution [19] has been widely used to describe the probability of the fault aging of transformer. Some researches considered the aging fault as the fault with independent operation condition. However, in practical situations, the aging fault usually depends on the thermal conditions of the transformer [20]. IEEE guidelines indicate that the insulation temperature is the control parameter of aging or insulation deterioration; therefore, the aging effect caused by HST is often considered in the aging studies. In 1984, Daijin's studies showed that the changing rate of physical properties of any measurable insulating material followed Arrhenius's reaction speed theory [21], which can be formulated in the form of the reaction rate as follows:

$$K_0 = A' \exp\left(\frac{B'}{\Theta_H + 273}\right) \tag{46.7}$$

In this formula, A' and B' are both the empirical constants.

Supposing that the insulation life of the transformer is proportional to the reciprocal of the reaction rate, Arrhenius's relative lifetime formula can be described by the following equation:

$$L = C \exp\left(\frac{B}{\Theta_H + 273}\right) \tag{46.8}$$

In Formula (46.8), L is the nominal insulation life, and B and C are empirical constants.

Constant B is recommended as 15,000, and C is related to the cooling mode; the specific calculation formula is shown as follows:

1. In the cooling mode of OA and FA:

C=0.0272*M (the weight of winding coil and core in kilograms) +0.01818*M (the weight of oil tank and accessories in kilograms) +5.034*M (the weight of oil in kilograms).

2. In the direct or indirect forced coil circulation cooling mode:

C=0.0272*M (the weight of winding coil and core in kilograms) +0.0272*M (the weight of oil tank and accessories in kilograms) +7.305*M (the weight of oil in kilograms).

When the Weibull distribution is applied to describe the aging process, the failure rate can be expressed by the following formula:

$$\lambda_w(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \tag{46.9}$$

In this formula, β is the shape factor, and η is the scale coefficient (typical life).

Make $\eta = L$, then the fault rate and cumulative probability distribution function of the Arrhenius–Weibull model can be expressed by the following formula:

$$\lambda_a(t\Theta_H) = \frac{\beta}{Ce^{\frac{B}{\Theta_H+273}}}\left(\frac{t}{Ce^{\frac{B}{\Theta_H+273}}}\right)^{\beta-1} \tag{46.10}$$

Note that two assumptions are made in this model. Firstly, even across different HSTs, the lifetime distribution has uniform shape parameters; secondly, for the whole or at least most transformers, during their insulation lifetime period, the initial parameters are unchangeable. According to these assumptions, on the basis of the data records of life-ending fault, adopt the maximum likelihood estimation (MLE) method to estimate the parameters.

46.4 System Structure and Function

In terms of the technical application, upon the introduction of the structure and function of software in Chap. 4, the software interface and the function of equipment management module interface as well as its application were introduced; after that, the equipment model analysis interface and its function were introduced with an actual example given afterward, and the number of 03-07-z-A001-0000000010 transformers was analyzed. Upon detecting and analyzing these HST data, its T- λ curve graph was given, in which we can clearly see that its current fault rate was 0.0002 to carry out a further analysis of the curve and detecting data. Focusing on its operation time T and fault rate λ, the HST model as built before was applied to carry out analysis and forecast on its lifetime evaluation, which can guide the maintenance plan. The management of the maintenance plan, while optimizing the maintenance process, has taken full account of the status of equipment, monitoring

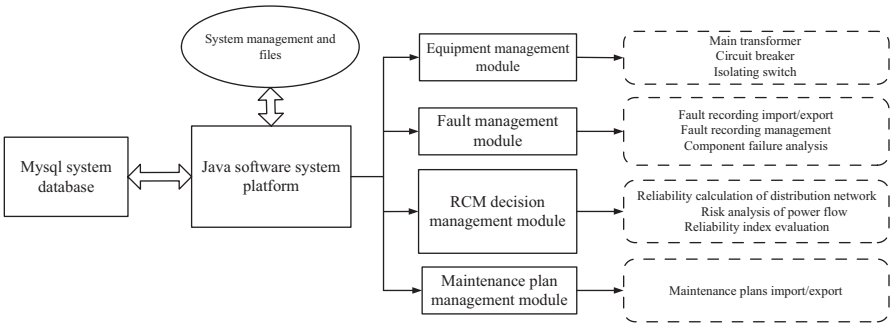
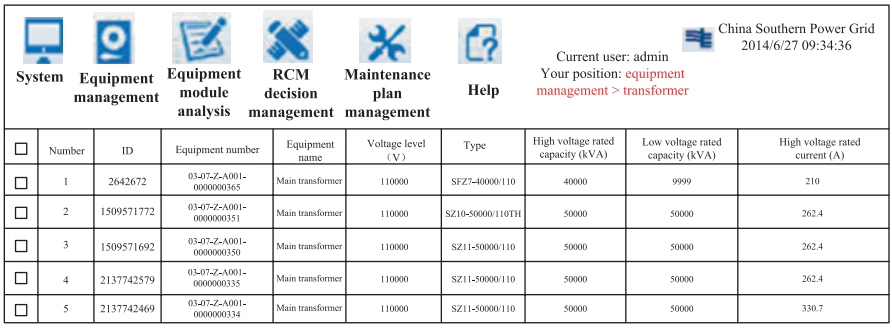


Fig. 46.2 System structure block diagram



	Number	ID	Equipment number	Equipment name	Voltage level (V)	Type	High voltage rated capacity (kVA)	Low voltage rated capacity (kVA)	High voltage rated current (A)
☐	1	2642672	03-07-Z-A001-0000000365	Main transformer	110000	SFZ7-40000/110	40000	9999	210
☐	2	1509571772	03-07-Z-A001-0000000351	Main transformer	110000	SZ10-50000/110TH	50000	50000	262.4
☐	3	1509571692	03-07-Z-A001-0000000350	Main transformer	110000	SZ11-50000/110	50000	50000	262.4
☐	4	2137742579	03-07-Z-A001-0000000335	Main transformer	110000	SZ11-50000/110	50000	50000	262.4
☐	5	2137742469	03-07-Z-A001-0000000334	Main transformer	110000	SZ11-50000/110	50000	50000	330.7

Fig. 46.3 Software interface of equipment management module

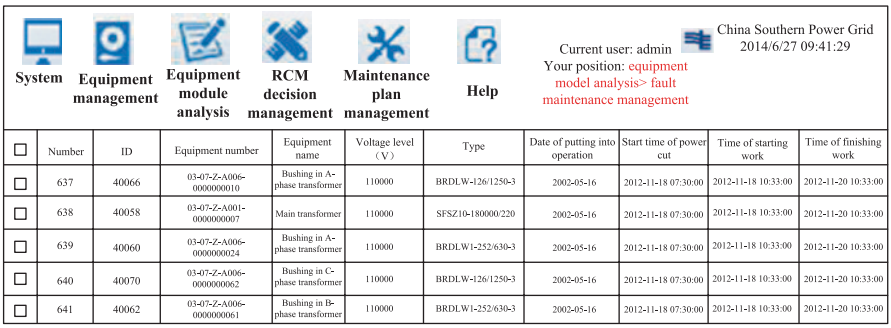
signal, maintenance history, testing data, reliability data, and so on, thus developing a scientific and reasonable maintenance scheme with attention paid to the weak points of maintenance, renovation, and replacement. According to the actual situation, efforts have been made to reduce unnecessary regular maintenance and prolong the cycle of large- and small-scale maintenance, thus to avoid the deficiency of problem because of lack of maintenance and also prevent over-maintenance.

Based on the aforementioned theory, the software analysis system is mainly developed by Java language while taking Mysql database as the background database. The transformer data from the management information system (MIS) are transmitted into the designed system, that is, the lifetime assessment and optimized maintenance system of the transformers based on HST model; then, the transformer data are calculated and analyzed in the software system. The designed system is divided into equipment management module, Reliability Calculation and Management (RCM) decision module, equipment model analysis module, and maintenance plan management module. The system structure block diagram is shown in Fig. 46.2.

According to Fig. 46.2, the equipment management module can query the basic data of the transformer from the background database, including equipment number, voltage level, capacity, and load rate; it can also revise data and implement, add, and delete operations, and store the data into the database.

The equipment management module software interface is shown in Fig. 46.3.

The current status of equipment, equipment monitoring data, the possible existing fault of equipment, etc. are mainly implemented by special personnel. They



	Number	ID	Equipment number	Equipment name	Voltage level (V)	Type	Date of putting into operation	Start time of power cut	Time of starting work	Time of finishing work
☐	637	40066	03-07-Z-A006-0000000010	Bushing in A-phase transformer	110000	BRDLW-126/1250-3	2002-05-16	2012-11-18 07:30:00	2012-11-18 10:33:00	2012-11-20 10:33:00
☐	638	40058	03-07-Z-A001-0000000007	Main transformer	110000	SFSZ10-180000/220	2002-05-16	2012-11-18 07:30:00	2012-11-18 10:33:00	2012-11-20 10:33:00
☐	639	40060	03-07-Z-A006-0000000024	Bushing in A-phase transformer	110000	BRDLW1-252/630-3	2002-05-16	2012-11-18 07:30:00	2012-11-18 10:33:00	2012-11-20 10:33:00
☐	640	40070	03-07-Z-A006-0000000062	Bushing in C-phase transformer	110000	BRDLW-126/1250-3	2002-05-16	2012-11-18 07:30:00	2012-11-18 10:33:00	2012-11-20 10:33:00
☐	641	40062	03-07-Z-A006-0000000061	Bushing in B-phase transformer	110000	BRDLW1-252/630-3	2002-05-16	2012-11-18 07:30:00	2012-11-18 10:33:00	2012-11-20 10:33:00

Fig. 46.4 Software interface of equipment fault maintenance management

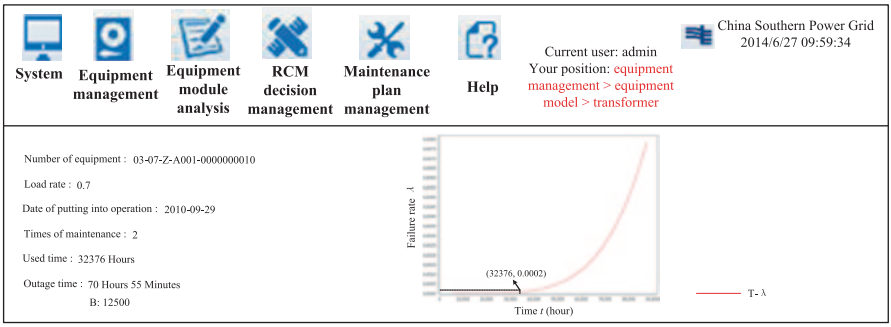


Fig. 46.5 *T*–*λ* curve interface diagram of transformer no. 03-07-z-A001-0000000010

record the monitoring condition of equipment, problems existing in equipment operation, abnormal situations found by operating personnel, etc. The maintenance personnel is responsible for the regular inspection, regular testing, the recording of calibration results, and analysis of evaluation results.

The equipment model analysis module is divided into fault records management module, device model, and fault analysis. The maintenance record history of the transformer can be queried in the fault records management, including the maintenance content and time, the time of power cuts, the fault records management, the equipment fault frequency, the fault position of equipment, and the effect on the reliability of the system, which are shown in Fig. 46.4.

The equipment model is based on the HST model and used to carry out assessment and analysis of the lifetime of transformers, calculate its real-time failure rate, evaluate the status of the transformer, and give suggestions on the maintenance. The software interface is shown in Fig. 46.4, in which the number of the transformer, its load rate, and date of putting into operation will be imported from the system's database automatically; meanwhile, its running time, and number and time of maintenance will be calculated by the backstage.

According to the lifetime assessment of the selected transformer by the HST model, its failure rate curve can be obtained, as shown in Fig. 46.5. As seen from Fig. 46.5, transformer no. 03-07-z-A001-0000000010 was put into operation on

September 29, 2010 with a load rate of 0.7, two maintenances, a running time of 32,376 , and an outage time of 70 h 55 min. From its T–λ curve, its current failure rate was 0.0002 (see the point (32376, 0.0002)).

The running time T and the failure rate λ curve T–λ can be used to analyze and forecast the lifetime of the transformer, and thus guide the maintenance plan. The fault analysis is carried on the correlation analysis to transformer maintenance record, which can get the fault analysis of main transformer manufacturers and main transformer faults of different types of transformers. Focusing on the equipment maintenance situation, knowing the maintenance information and comparing it with the reliability database, we can find which manufacturer's transformer is of good or bad quality. It can be used not only for maintenance management but also for guiding the equipment procurement and invitation of tender.

As to the maintenance plan management, firstly, based on the service life of the equipment situation, historical maintenance circumstances, operation reliability data of the system, etc. determine the system to be analyzed; secondly, analyze the existing defects and possible faults, and analyze what effect these faults will bring to the system function and equipment operational reliability; then, develop various methods to eliminate defects; for example, the system operation instability fault is mainly the faulty design or the equipment quality, or the poor installation and maintenance, or the lack of operation experience; with attention to different reasons and faults, different processing methods are used. When the maintenance schemes are put forward, a comprehensive analysis of the economic efficiency of various schemes has to be carried out, for example, whether simple maintenance or the replacement of a single element should be selected. If the device status of the whole system is very poor and it will be eliminated sooner or later, the replacement of a single element is not economical; in contrast, if the operation maintenance is very good because the system has been put into operation, and the system can prolong the service life upon analysis only with individual equipment aging frequently, it would be better to replace the element.

46.5 Optimized Maintenance Mode of Transformers

Based on the aforementioned description, the purpose of equipment maintenance of a power system is to improve the system operation reliability and the operation rate of equipment. As to an enterprise, improvement of the economic efficiency is the ultimate goal; in this sense, the improvement of equipment reliability should be implemented surrounding the improvement of economic benefit of the enterprise; therefore, whether the optimization of the maintenance mode of the transformer can improve the reliability of the equipment and economic efficiency of enterprises is the current standard method to evaluate the system and its reliability analysis, which is also the current study status.

As the optimization of the maintenance mode decision process has fully considered the equipment situation, monitoring signal, historical maintenance, testing

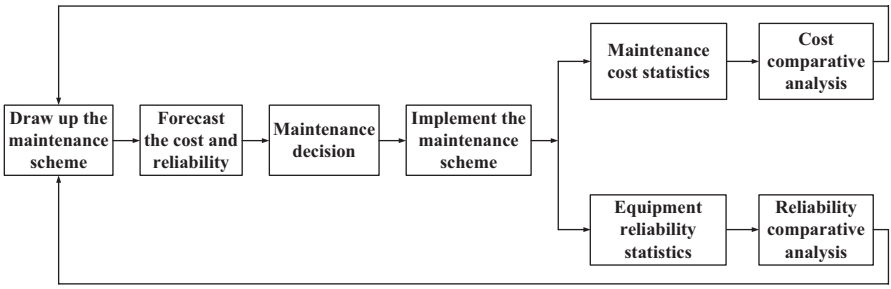


Fig. 46.6 System flow chart of lifetime assessment and maintenance optimization of a transformer

data, reliability data, etc. the developed maintenance scheme is more scientific and reasonable and can be targeted to strengthen the weak part of the maintenance and renovation until the update of the generation. Based on the actual situation, efforts can be made to reduce unnecessary regular maintenance or prolong the period of large- and small-scale maintenance so as to avoid the deficiency of problem because of lack of maintenance and prevent over-maintenance.

We have established a transformer lifetime assessment model and a system flow chart of maintenance optimization to evaluate the optimized maintenance mode based on the reliability and the maintenance cost as examination indexes, which is shown in Fig. 46.6.

As seen from the Fig. 46.6, the optimization of the maintenance mode has an obvious effect on reducing the enterprise's production cost and improving the economic efficiency of the enterprise. Setting the transformer maintenance as an example, the transformer is of 500 MVA, which costs several millions to tens of millions Chinese Ren Min Bi (RMB) for each maintenance; in this sense, the reduction of one time maintenance can save millions or even tens of millions RMB. In addition, the maintenance reduction can improve the utilization coefficient of the transformer and the equipment reliability.

In terms of different types of power transmission and transformation equipment, the lifetime evaluation and optimized maintenance system of a transformer based on the HST model can implement HST detection on different types of power equipment accordingly; then, the detected data will be input into the software according to the corresponding equipment lifetime evaluation model with the load rate computed and the lifetime evaluation analyzed. When the data have been put into operation, the software in the background computer will compute its entire operating time, maintenance frequency, and time; therefore, the lifetime evaluating and optimized maintenance system designed in this chapter can not only evaluate the lifetime of power transformer but also play a role in carrying out detailed lifetime risk evaluation and operation analysis of different types of power transmission and transformation equipment as well as other power monitoring systems.

46.6 Conclusion

The lifetime assessment and optimized maintenance system of a transformer based on the HST model can overcome the application difficulties of power transmission and transformation equipment in state maintenance work. It can play a very active role in the transformer maintenance; besides, the application of the system designed in this chapter can effectively improve the reliability of power grid equipment, reduce the production cost, and enhance the profit ability of power enterprises as well.

References

1. Shi Y, Liu Z. Effective way of raising reliability level of oil temperature meter for transformer. Xuchang, Electr Power Constr. 2000;22(7):64–6 (In Chinese).
2. Zhang D. Maintenance and replacement strategies of electric power equipment based on condition monitoring and system risk evaluation. Chongqing: Chongqing University; 2012.
3. Su R, Han B, Lu Y, et al. Transformer's economic life assessment based on random and fuzzy theory. Xuchang, Power Syst Prot Control. 2014;42(1):9–16 (In Chinese).
4. Diao Y. Reliability analysis of large power transformers. Beijing, Trans China Electrotech Soc. 1990;5(4):43–8 (In Chinese).
5. Zhang A. Application of fault tree analysis to reliability study of electric power system. Shanghai, East China Electr Power. 2005;33(2):14–7 (In Chinese).
6. Hu Y. The fault diagnose of electric power transformer basing on fault tree analysis. Hefei: HeFei University of Technology; 2002.
7. Duan J, Li H. Based on fault tree's failure diagnosis expert system research. Beijing, Sci Technol Eng. 2009;9(7):1914–7 (In Chinese).
8. Zhang C, Liu W. Fault diagnosis expert system based on fault tree. Mianyang, Ordnance Ind Autom. 2009;28(11):15–6, 23 (In Chinese).
9. Liao R, Xiao Z, Gong J, et al. Markov model for reliability assessment of power transformers. Wuhan, High Volt Eng. 2010;36(2):322–8 (In Chinese).
10. Jiang X. Maintenance optimization for power equipment based on Markov process. Jinan: Shandong University; 2013.
11. Ding M, Li S, Liu S. A new algorithm for calculating frequency and duration indices in reliability analysis of composite generation and transmission system. Nanjing, Autom Electr Power Syst. 1999;23(22):12–5, 62 (In Chinese).
12. Zhou N, Ma J, Hu B, et al. Reliability tracing technique for power transformers using the fault tree analysis method. Xuchang, Power Syst Prot Control. 2012;40(19):72–7 (In Chinese).
13. Xing R, Yao A, Xie K, et al. Power flow tracing method for impact analysis of large scale power grid's reliability. Beijing, Power Syst Technol. 2006;30(10):54–8 (In Chinese).
14. Mora-Florez J, Bedoya-Hernandez J, Perez-Londono S. Fault tree based methodology to evaluate the reliability of converter transformers. Transmission and Distribution conference and Exposition, IEEE/PES; 2008. p. 1–7.
15. Hu B, Xie K, Li X, et al. Techniques of tracing the unreliability contributions of HVDC transmission system components. Beijing, Proc CSEE. 2000;23(10):29–35 (In Chinese).
16. Diao Y. Reliability analysis of large power transformers. Beijing, Trans China Electrotech Soc. 1990;5(4):43–8 (In Chinese).
17. Xie K, Billinton R. Tracing the unreliability and recognizing the major unreliability contribution of network components. Engl, Reliab Eng Syst Saf. 2009;94(5):927–31.
18. Lu X, Lu S. Reliability assessment based on the kth ordered experiment life of Weibull distribution. Shanghai, Mach Des Res. 2005;21(6):10–2 (In Chinese).

19. Volkanovski A, Čepin M, Mavko B. Application of the fault tree analysis for assessment of power system reliability. USA, Reliab Eng Syst Saf. 2009;94(6):1116–27.
20. Chen J. Analysis of thermal aging problem of cast-resin dry-type transformer. Shenyang, Transformer. 2002;49(3):14–6 (In Chinese).
21. Dhillon BS. Reliability, quality, and safety for engineers. Boca Raton: CRC Press; 2005. p. 25–8.

Chapter 47
Transmission Line Parameter Coupling Algorithm Based on Synchronous Sampling Signals

Le Bao and Wei Cong

Abstract The existing online line parameter calculation methods are widely considered complicated, inaccurate, and incomplete. Thus, a coupling calculation method of transmission parameters which is based on synchronous sampling signals at two terminals is proposed in this chapter. This chapter adopts the π-type lumped parameter model of transmission line. Based on voltage and current information at the two terminals of the transmission line, the coupling algorithm is realized by solving the differential circuit equations describing the relationship of line parameters. This chapter uses single-circuit and double-circuit lines under asymmetry state as example and presents the calculating process. This method has simple principles and an easy realization, and obtains the transmission line parameters accurately and completely. Finally, power systems computer aided design (PSCAD) simulation results verify effectiveness and correctness.

Keywords Synchronous sampling signal · Transmission line parameters · π-type lumped parameter model · Coupling calculation · Online calculation

47.1 Introduction

An overhead transmission line is the carrier of electric power, which is always an important part of the power system. Inaccurate transmission line parameters have an effect on the stabilization of the power system. Since it can improve accuracy of state estimation, power flow calculation, fault analysis, and relay protection setting calculation, it is necessary and significant to acquire the parameters exactly [1–2].

Online calculation methods based on the phasor measurement unit (PMU) mainly use synchronized phasor to calculate transmission line parameters. Actually, considering the high cost and the limitation of data sharing, this online calculation method only applies at a small scale. Moreover, based on synchronous

L. Bao (✉) · W. Cong
Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University), Ministry of Education, 250061 Jinan, China
e-mail: baomingle@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_47

sampling signals at two terminals, some other online calculation methods can calculate the transmission line parameters as well by solving integral or differential equations. However, at home and abroad, this kind of approach is based on the "one-type" lumped parameter model, which ignored the admittance against the ground and the admittance between different transmission lines [3–4]. So, these algorithms are incomplete and incorrect. Based on synchronous sampling signal, the online coupling algorithm of transmission line parameters is presented in this chapter.

47.2 The Principle of the Online Calculation Method

In practice, the lumped parameter model is used more frequently for overhead transmission lines whose length is less than 300 km, such as "pi-type," "one-type," and "T-type." Taking the "pi-type" model of a single-phase system for example, the synchronous sampling signal-based calculation method of line parameters is introduced and generalized to a three-phase system. Figure 47.1 shows the pi-type lumped parameter model for a single-phase system.

In Fig. 47.1, u_m, u_n, i_m, and i_n are the instantaneous values of voltage and current at terminal M and terminal N of a transmission line; *R, X,* and *C* are the equivalent resistance, inductive reactance, and distributed capacitance, respectively; i_{c1}, i_{c2} are instantaneous values of the current passing through the distributed capacitance.

According to Kirchhoff's Current Law (KCL) and Kirchhoff's Voltage Law (KVL), the differential equations of a transmission line are:

$$i_m - i_n = i_{c1} + i_{c2} = \frac{C}{2}\frac{du_m}{dt} + \frac{C}{2}\frac{du_n}{dt} \tag{47.1}$$

$$u_m - u_n = R(i_m - i_{c1}) + L\frac{d(i_m - i_{c1})}{dt} \tag{47.2}$$

Further, the discrete differential equations are:

$$i_m(k) - i_n(k) = \frac{C}{2}\frac{u_m(k+1) - u_m(k-1)}{2T} + \frac{C}{2}\frac{u_n(k+1) - u_n(k-1)}{2T} \tag{47.3}$$

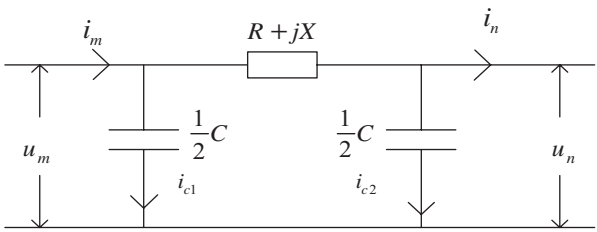


Fig. 47.1 π-type lumped parameter model for transmission line

$$u_m(k)-u_n(k)=R(i_m(k)-i_{c1}(k))+L\frac{i_m(k+1)-i_{c1}(k+1)-i_m(k-1)+i_{c1}(k-1)}{2T} \tag{47.4}$$

While several groups of sampling signals at terminal M and terminal N are brought into Eqs. 47.3 and 47.4, the value of the line parameters (*R, X,* and *C*) could be solved out through the least square method. This is the principle of online calculation methods of line parameters in this chapter [1].

47.3 The Coupling Algorithm Applications in a Three-Phase System

In practice, when the three-phase system works under asymmetry state, the zero sequence parameters of the transmission line could be calculated out. In this method, the synchronous sampling signals at both ends of the transmission line are adopted as input data. Then the mathematical relationship of different line parameters in the pi-type model is described directly by differential equations. Finally, sequence parameters of the transmission line can be calculated through solving these equations. Figure 47.2 shows a pi-type lumped parameter model of a three-phase system.

In Fig. 47.2, C_s and C_m are the distribution capacitance against the ground and between two-phase circuits. R_s, L_s, R_m, and L_m are the self-and mutual impedances of the one-phase circuit, respectively.

Based on KCL, the differential equations of phase A are obtained as follows:

$$\begin{cases} i_{mA}-i_{nA}=i_{csA}+i_{cmAB}+i_{cmAC} \\ i_{csA}=i_{csAm}+i_{csAn}=\dfrac{C_s}{2}\dfrac{du_{mA}}{dt}+\dfrac{C_s}{2}\dfrac{du_{nA}}{dt} \\ i_{cmAB}=i_{cmABm}+i_{cmABn}=\dfrac{C_m}{2}\dfrac{d(u_{mA}-u_{mB})}{dt}+\dfrac{C_m}{2}\dfrac{d(u_{nA}-u_{nB})}{dt} \\ i_{cmAC}=i_{cmACm}+i_{cmACn}=\dfrac{C_m}{2}\dfrac{d(u_{mA}-u_{mC})}{dt}+\dfrac{C_m}{2}\dfrac{d(u_{nA}-u_{nC})}{dt} \end{cases} \tag{47.5}$$

Take several groups of sampling signals at terminal M and terminal N into Eq. 47.5; the value of C_s and C_m could be solved out through the least square method. According to KVL, the differential equation is:

$$u_{mA}-u_{nA}=R_s i_A+R_m(i_B+i_C)+L_s\frac{di_A}{dt}+L_m\left(\frac{di_B}{dt}+\frac{di_C}{dt}\right) \tag{47.6}$$

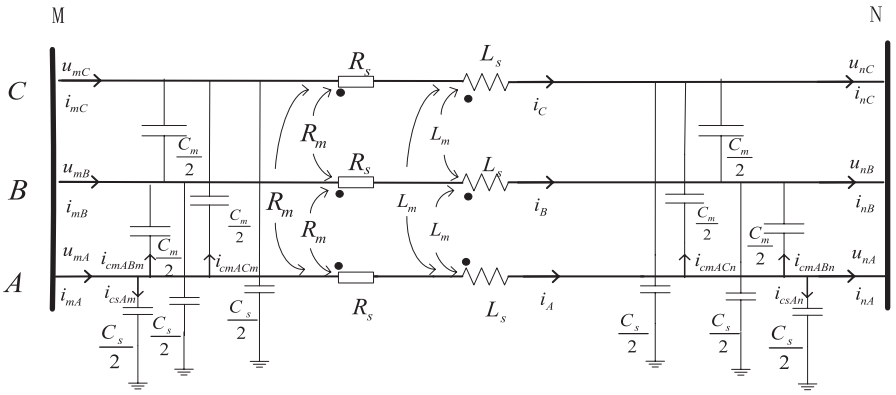


Fig. 47.2 π-type lumped parameter model of a three-phase system

Assume that *n* groups of sampling signals at terminal M and terminal N have been carried into Eq. 47.6, and then we can get the matrix equation.

Because the maximum rank of the coefficient matrix is less than the number of unknown variations, only using the data of phase A is not enough. As the three-phase system works under asymmetry state, we can list the differential equations of phase B and C on the base of KVL. Then the phase parameters (R_s, R_m, L_s, and L_m) can be obtained by using the least square method. Eq. 47.7 shows the transition between phase parameters and sequence parameters of a transmission line in a three-phase system. Finally, it is possible to get the positive and zero sequence parameters of the transmission line.

$$\begin{cases} Z_1 = Z_s - Z_m & C_1 = C_s + 3C_m \\ Z_0 = Z_s + 2Z_m & C_0 = C_s \end{cases} \tag{47.7}$$

47.4 The Coupling Algorithm Applications in aDouble Circuit

The advantages of double-circuit transmission lines on the same tower are occupying little space and having large capacity in transmission. So, it is widely applied in high-voltage power lines. Compared with single-circuit lines, mutual parameters between double circuits should be taken into account in the line parameters, which are very important for analyzing the features of double-circuit lines [5].

Likewise, present methods such as incremental approach and differential means cannot get complete parameters of transmission lines. Here, we use double circuits with the same voltage class but from different buses as example to illustrate the

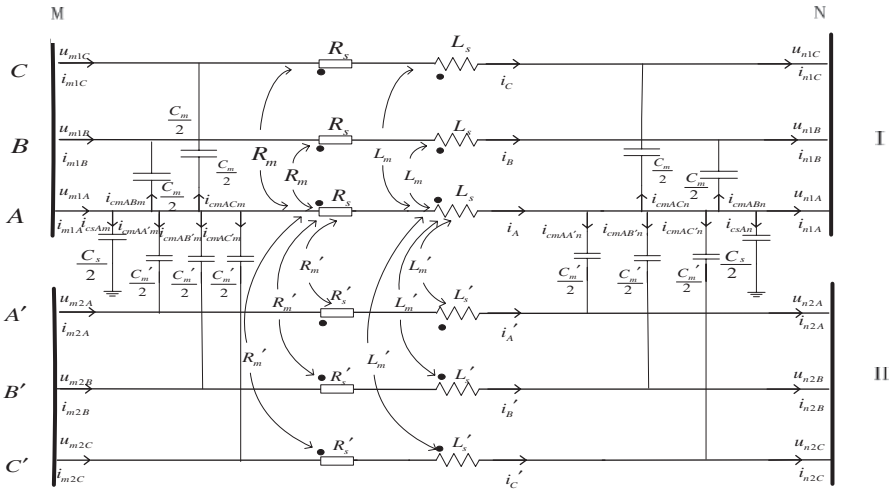


Fig. 47.3 π-type lumped parameter model of phase A in double circuit installed on same tower

coupling algorithm of mutual parameters. Figure 47.3 shows a pi-type lumped parameter model of phase A in a double circuit installed on the same tower. C_s, C_m, and C_m' are the distribution capacitance of phase A against the ground, between different phases, and between different circuits; and R_s, L_s, R_m, L_m, R_m', and L_m' are the self- and mutual impedance, respectively. The differential equations for phase A can be established on the base of KCL and KVL as shown in Eqs. 47.8 and 47.9.

$$\begin{cases} i_{m1A} - i_{n1A} = i_{csA} + i_{cmAB} + i_{cmAC} + i_{cmAA'} + i_{cmAB'} + i_{cmAC'} \\ i_{csA} = i_{csAm} + i_{csAn} = \dfrac{C_s}{2}\dfrac{du_{m1A}}{dt} + \dfrac{C_s}{2}\dfrac{du_{n1A}}{dt} \\ i_{cmAB} = i_{cmABm} + i_{cmABn} = \dfrac{C_m}{2}\dfrac{d(u_{m1A} - u_{m1B})}{dt} + \dfrac{C_m}{2}\dfrac{d(u_{n1A} - u_{n1B})}{dt} \\ i_{cmAC} = i_{cmACm} + i_{cmACn} = \dfrac{C_m}{2}\dfrac{d(u_{m1A} - u_{m1C})}{dt} + \dfrac{C_m}{2}\dfrac{d(u_{n1A} - u_{n1C})}{dt} \\ i_{cmAA'} = i_{cmAA'm} + i_{cmAA'n} = \dfrac{C_m'}{2}\dfrac{d(u_{m1A} - u_{m2A})}{dt} + \dfrac{C_m'}{2}\dfrac{d(u_{n1A} - u_{n2A})}{dt} \\ i_{cmAB'} = i_{cmAB'm} + i_{cmAB'n} = \dfrac{C_m'}{2}\dfrac{d(u_{m1A} - u_{m2B})}{dt} + \dfrac{C_m'}{2}\dfrac{d(u_{n1A} - u_{n2B})}{dt} \\ i_{cmAC'} = i_{cmAC'm} + i_{cmAC'n} = \dfrac{C_m'}{2}\dfrac{d(u_{m1A} - u_{m2C})}{dt} + \dfrac{C_m'}{2}\dfrac{d(u_{n1A} - u_{n2C})}{dt} \end{cases} \quad (47.8)$$

$$u_{m1A} - u_{n1A} = R_s i_A + R_m (i_B + i_C) + R_m{}' (i_{A'} + i_{B'} + i_{C'}) + L_s \frac{di_A}{dt} + L_m \left(\frac{di_B}{dt} + \frac{di_C}{dt} \right) + L_m{}' \left(\frac{di_{A'}}{dt} + \frac{di_{B'}}{dt} + \frac{di_{C'}}{dt} \right) \tag{47.9}$$

Follow the same steps as above to list the differential equations of phase B and C, and then the phase parameters can be obtained by using the least square method. Finally, the sequence parameters can be obtained by bringing them into equation of phase sequence conversion.

This coupling algorithm is technically feasible for the engineering application. In fact, the sampling information can be gathered conveniently in protections and measure-control devices which have unified time. For the reasons of unbalance load and external disturbances, the power system often works under asymmetry state. Then we can calculate the transmission line parameters by the coupling algorithm based on these sampling signals.

47.5 Computer Simulations

In this chapter, we build two models of three-phase systems and double-circuit transmission lines using the power systems computer aided design (PSCAD)/ Electro Magnetic Transient in Direct Current System (EMTDC) software. Transmission lines are built in the Bergeron model and their lengths are 50, 100, and 200 km. Table 47.1 shows transmission line parameters in unit length.

47.5.1 Applications in a Three-Phase System

In order to get the zero sequence parameters, we assume that the system was waiting for single-phase reclosing. For a length of the transmission line of 50 km, results of the coupling algorithm are shown in Fig. 47.4.

As shown in Fig. 47.4, all the parameters are stable and have no fluctuation. Table 47.2 shows the relative error of coupling algorithm and decoupling method for different lengths of transmission lines.

Table 47.1 Table of transmission line parameters in simulation

Parameters	Single circuit		Double circuit		
	+ve SEq.	0 SEq.	+ve SEq.	0 SEq.	0 SEq. mutual
R (Ω/km)	0.0116	0.6535	0.022	0.1828	0.1608
X (Ω/km)	0.3139	0.9942	0.28	0.86	0.58
C (μF/km)	0.01081	0.0025	0.0132	0.0088	0.0025

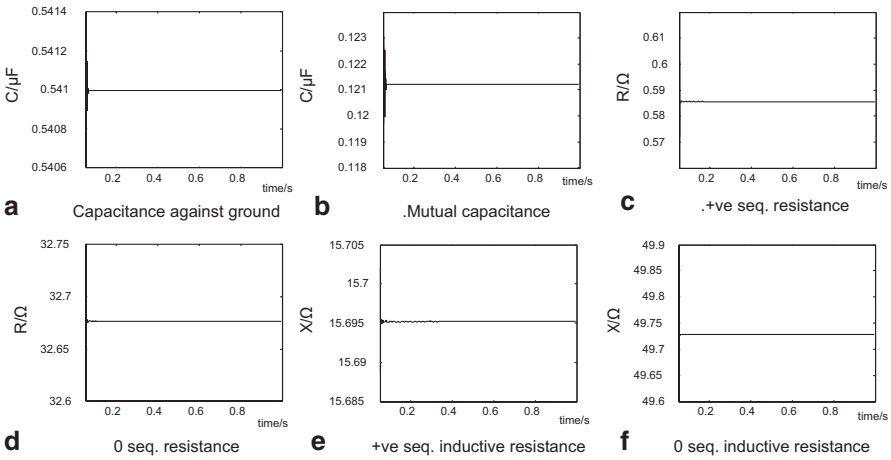


Fig. 47.4 Simulation results of transmission line parameters

Table 47.2 Simulation results of transmission line parameters of single circuits

Length	Method	Parameters	$R_+(\Omega)$	$R_0(\Omega)$	$X_+(\Omega)$	$X_0(\Omega)$	$C_+(\mu F)$	$C_0(\propto F)$
50 km	Decoupling	Results	0.64	32.73	15.67	49.75	–	–
		Error%	9.98	0.19	0.15	0.09	–	–
	Coupling	Results	0.585	32.67	15.69	49.72	0.541	0.121
		Error%	0.96	0.005	0.001	0.036	0.09	1.3
100 km	Decoupling	Results	1.577	65.71	31.15	99.63	–	–
		Error%	35.9	0.54	0.75	0.2	–	–
	Coupling	Results	1.161	65.20	31.34	99.40	1.082	0.243
		Error%	0.13	0.23	0.13	0.012	0.16	0.79
200 km	Decoupling	Results	5.5141	133.3	60.65	200.1	–	–
		Error%	137	1.97	3.4	0.66	–	–
	Coupling	Results	2.28	129.2	62.36	198.4	2.171	0.489
		Error%	1.65	1.11	0.66	0.21	0.43	0.39

47.5.2 Applications in Double-Circuit Lines

For the length of a transmission line of 100 km, results of the coupling algorithm are shown in Fig. 47.5.

Make a comparison between coupling algorithm and decoupling method for different lengths of transmission lines.The result is shown in Table 47.3.

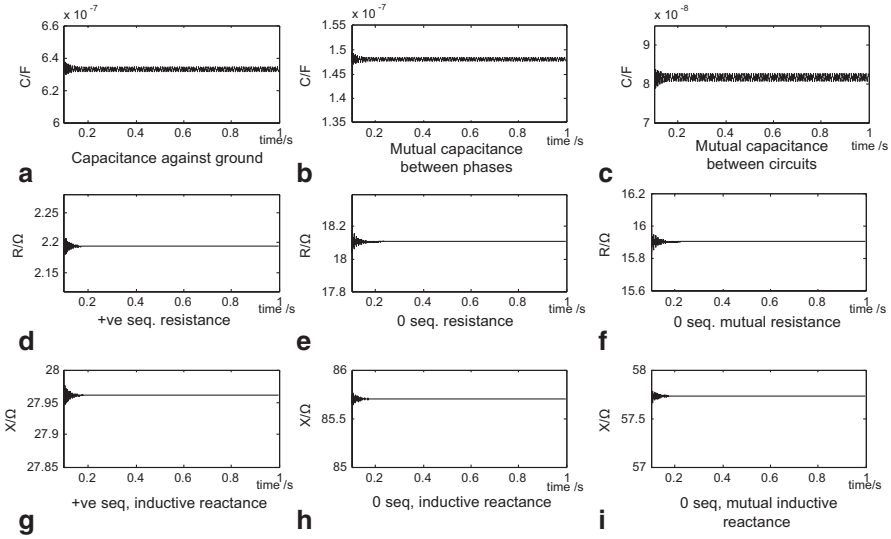


Fig. 47.5 Simulation results of parameters in a double-circuit system

Obviously, compared with the decoupling method, no matter whether in a three-phase system or in double-circuit lines, the parameters calculated by the coupling algorithm are more complete and the precision is much higher.

47.6 Conclusion

This chapter presents an online coupling algorithm of transmission line parameters. Based on the pi-type lumped parameter model, it establishes the differential equations describing the relationship of line parameters. Then this method adopts synchronous sampling signals at two terminals of the transmission line to solve the system of differential equations by means of the least square method. The advantage of the coupling algorithm is that it can get the transmission line parameters accurately and completely. Because the disadvantage of using the lumped parameter model is that it is not applicable for long-distance transmission lines.

Table 47.3 Simulation results of transmission line parameters in double-circuit system

Length (km)	Method	Parameters	$R_+(\Omega)$	$R_0(\Omega)$	$R_m(\Omega)$	$X_+(\Omega)$	$X_0(\Omega)$	$X_m(\Omega)$	$C_+(\mu F)$	$C_0(\mu F)$	$C_m(\mu F)$
50	Decoupling	Results	1.18	9.19	8.09	14.1	43.1	29.1	–	–	–
		Error (%)	7.58	0.59	0.59	0.3	0.25	0.31	–	–	–
	Coupling	Results	1.11	9.13	8.03	14	42.9	28.9	0.66	0.4	0.13
		Error (%)	0.6	0.09	0.15	0.01	0.05	0.07	0.01	0.08	2.6
100	Decoupling	Results	2.81	18.6	16.4	28.3	86.7	58.6	–	–	–
		Error (%)	27.8	1.82	1.93	1.15	0.87	1.1	–	–	–
	Coupling	Results	2.19	18.1	15.9	27.9	85.7	57.7	1.32	0.87	0.24
		Error (%)	0.26	0.95	1.08	0.14	0.35	0.46	0.07	0.19	2.0
200	Decoupling	Results	9.36	39.3	34.7	58.4	178	121	–	–	–
		Error (%)	112	7.39	8.03	4.25	3.41	4.38	–	–	–
	Coupling	Results	4.33	35.1	30.7	55.6	169	114	2.65	1.73	0.48
		Error (%)	1.59	4.15	4.56	0.72	1.62	2.15	0.37	1.36	4.6

References

1. Hu ZJ, Chen YP, Xu W, et al. Principles and realization of live line measurement of parameters of transmission lines with mutual inductance based on differential equations. Proc CSEE. 2005;25(2):28–33 (In Chinese).
2. Tern A, Javidi B. Three-dimensional image sensing, visualization, and processing using integral imaging. Proc IEEE. 2006;94(3):591–607.
3. Li SS, Chen XY, Qi GQ, et al. Discussion on live line measurement of the parameters of the transmission lines with four-circuit on a tower. High Volt Eng. 2006;32(7):17–20 (In Chinese).
4. Chen KC, Damrau KM. Accuracy of approximate transmission line formulas for overhead wires. IEEE Trans Electromagn Compat. 1989;31(4):396–7.
5. Suonan JL, Zhang YN, Qi J. Study of current differential protection using time-domain capacitive current compensating algorithm on pi-model. Proc CSEE. 2006;26(5):12–8 (In Chinese).

Chapter 48
Analysis of Electromagnetic Interference Effects on Gas Pipelines Due to a Nearby Parallel UHV Transmission Line

Jingzhuo Zhang, Xishan Wen, Wei Li, Hailiang Lu and Yu Liu

Abstract The inductive coupling between ultra-high voltage (UHV) transmission lines and nearby underground pipelines raises the potential of pipelines which threatens the pipeline security and reliability. In this chapter, a transmission line–pipeline model is established based on the commercial software CDEGS, and the electromagnetic interference is simulated and analyzed. The pipeline potentials at different parallel distances are calculated in consideration of the effect of the overhead grounding wire. The simulation results are in conformity with the theoretical analysis. On this basis, a real 750-kV transmission line, and the nearby gas line model are built and simulated. The results show that the long distance of parallel and oblique approach can lead the maximum pipeline grounding potential raise (GPR) to exceed 100 V, i.e., 10 times of the safety threshold of 10 V. The solution is proposed and simulated in this chapter, and the maximum GPR of a pipeline is reduced to less than 10 V.

Keywords Electromagnetic interference · Induced AC voltages · Gas pipeline · Transmission line · Mitigation methods

48.1 Introduction

When the transmission lines and pipelines share the same corridor, attention shall be paid to the electromagnetic interference effects on the pipelines [1]. Studies of electromagnetic interference effects on gas pipelines due to nearby parallel ultra-high voltage (UHV) transmission lines show that the assessment criteria on the extent of the electromagnetic interference effects are lacking in China. Safe distance

J. Zhang (✉) · X. Wen · W. Li · H. Lu · Y. Liu
College of Electronic Engineering, Wuhan University, 430072 Wuhan, China
e-mail: zhangjingzhuo@outlook.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_48

between transmission lines and pipelines needs to be determined immediately to guide the route planning in engineering design [2, 3]. When the path of the transmission lines and pipelines cannot be changed due to the corridor limits, effective mitigation methods have to be applied [4, 5]; therefore, it is crucial to analyze the electromagnetic interference effects on pipelines and decide the assessment criteria. As to the safety considerations, effective mitigation methods should be studied and applied.

In this chapter, we present a study to determine the electromagnetic interference effects on gas pipelines which are parallel to a 750-kV transmission line constructed nearby. The study consists of the following content: (1) a transmission line–pipeline model which calculates the potentials on the gas pipelines in different distances between the transmission line and the pipeline, (2) computation of the gas pipeline grounding potential raise (GPR) due to a nearby 750-kV transmission line, and (3) design of a mitigation system which reduces the interference level and analysis of the simulation results.

48.2 Calculation Models

Determination of the interference effects in a typical right-of-way is a complex procedure. It not only requires a good knowledge of the transmission line and the gas pipeline layouts as well as the electrical characteristics and parameters but also an accurate representation of the soil structure; therefore, both the soil model and the transmission line–pipeline model are established to study the pattern of the induced alternating current (AC) voltage.

48.2.1 Soil Model

The test results of soil resistivity along the joint corridor of the transmission line and the gas pipeline are shown in Table 48.1.

According to the results presented in Table 48.1, more stringent conditions are taken into account. We take 150Ωm as a uniform soil resistivity in the following calculations.

Table 48.1 Soil resistivity along the joint corridor of the transmission line and the gas pipeline at 2 and 4 m (Ωm)

Number of measuring position	1	2	3	4	5	6	7	8	9
Soil resistivity at 2 m	8.3	24.9	2.6	20.7	19.4	27.6	13.8	19.4	69.1
Soil resistivity at 4 m	3.3	36.1	2.8	19.6	95.6	31.5	17.2	21.9	41.9

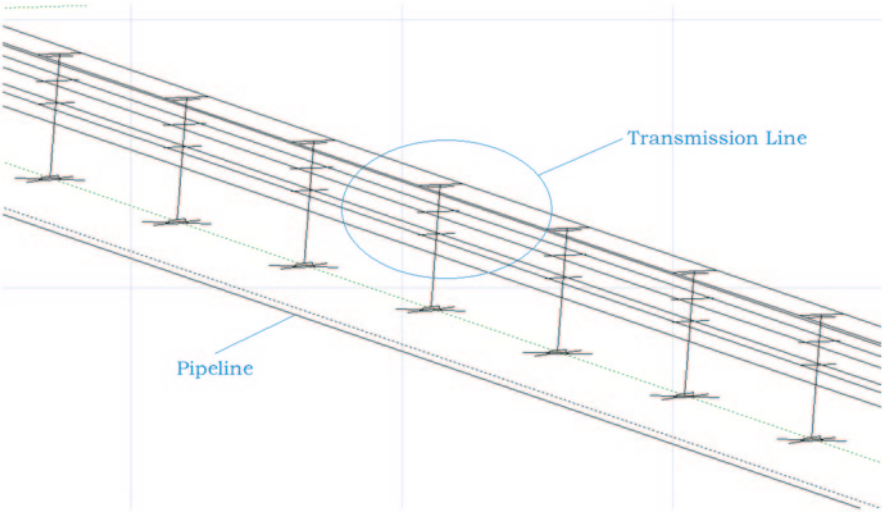


Fig. 48.1 Transmission line–pipeline model

48.2.2 Transmission Line–Pipeline Model

The existing transmission line is a 750–kV-double-circuit line in inverse phase sequence with a nominal height of 49.45 m and a line current of 1500 A. All the simulations in this chapter assume steady-state conditions. The diameter of the gas pipe is 610 mm, the thickness is 10 mm, the buried depth is 1.2 m, the side length of the tower grounding grid is 15 m, the ray length is 20 m, and the grounding resistance is 2.6 Ω in the given soil resistivity. The transmission line–pipeline model is shown in Fig. 48.1.

48.3 Simulation Results and Analysis

48.3.1 Simulation of the Transmission Line–Pipeline Model

Upon changing the distance between the transmission line and the pipeline, the potentials on the pipeline are calculated with the pipeline parallel to the transmission line along 10 km under ideal conditions. The results are shown in Fig. 48.2.

From the calculation results, potentials at both ends of the pipe are the largest, while potentials in the middle of the pipe are the smallest. The inductive voltage on the pipe decreases rapidly with the increase in the distance between pipeline and transmission line. The distribution characteristics of the pipeline potentials match the theoretical analysis of the transmission line–pipeline model. According to the national standard SY/T0087.1–2006 *the Evaluation Standard of Steel Pipe and Storage Tank Corrosion-The Direct Assessment of Buried Steel Pipeline External*

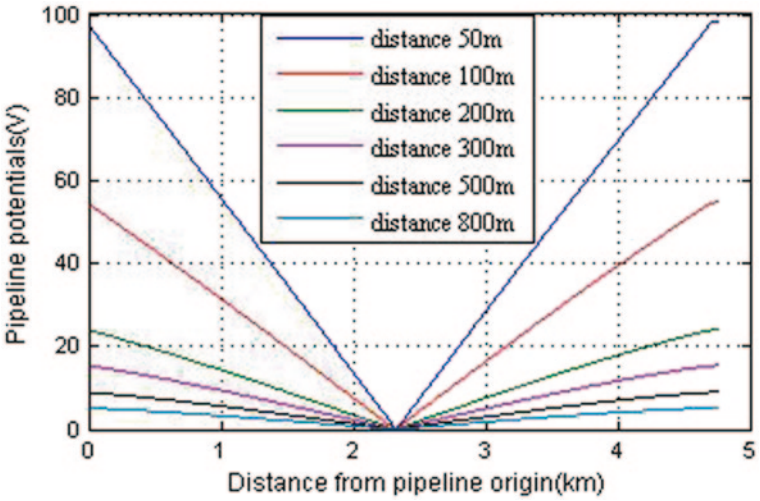


Fig. 48.2 Potentials on the pipeline when the distance between the pipeline and the transmission line is changing

Corrosion [6], the potentials on the pipe should be controlled below 10 V due to the alkaline soil; therefore, in order to meet the requirement, the maximum pipeline potential should be less than 10 V. The variation trend of the maximum pipeline potentials is shown in Fig. 48.3, while the distance between the pipeline and the transmission line is increasing.

As shown in Fig. 48.3, maximum pipeline potentials decrease rapidly with the increase in the distance between the pipeline and the transmission line, but the rate of change decreases gradually. When the distance is above 477.33 m, it can be

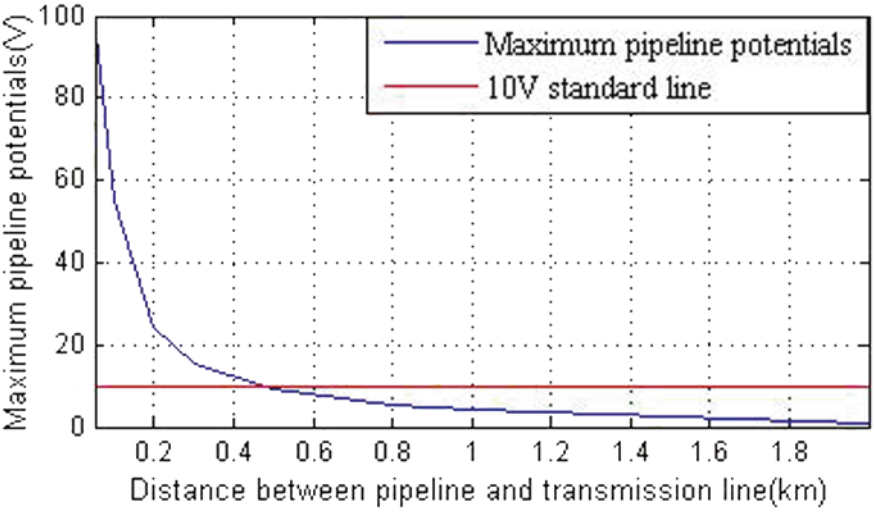


Fig. 48.3 Relationship between maximum pipeline potentials and transmission line–pipeline distance

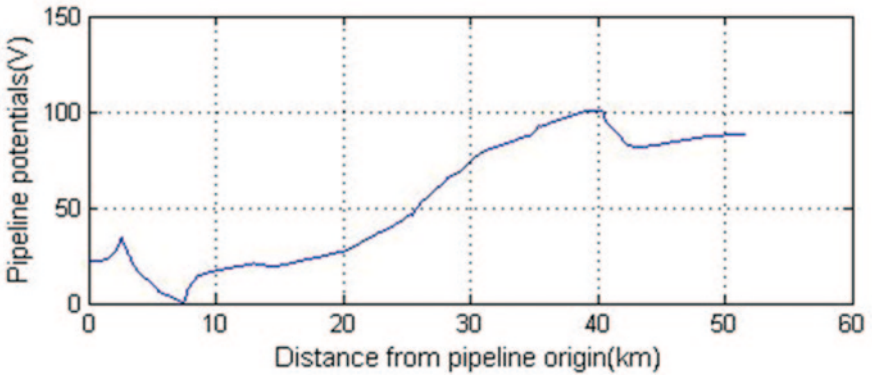


Fig. 48.4 Potentials along the gas pipeline

considered as no interference between the pipeline and the transmission line during normal operation.

When the transmission line and pipeline are vertically crossed, the maximum potential on the pipe is below 0.4 V according to the simulation result; therefore, there is no need to analyze the safe distance when transmission line and pipeline are vertically crossed.

48.3.2 Simulation of the Existing 750-kV Transmission Line–Gas Pipeline Model

The existing 750-kV transmission line is located in Ningxia Province, China. The capacity of the 750-kV transmission line is 2300 MVA; so the maximum load current is assumed to be 1500 A. The simulation result is shown in Fig. 48.4.

Figure 48.4 shows the gas pipeline GPR due to the presence of the nearby 750-kV transmission line. Due to the small cross angle and short distance between the gas pipeline and the transmission line, the induced AC voltages on the pipeline are quite high at most positions.

48.3.3 Mitigation Methods

As the locations of the transmission line and the pipeline are fixed, we need to install mitigation systems on pipelines to reduce the potential difference between the pipe and the ground. We provide two kinds of mitigation methods to solve the problem.

Mitigation Method 1: Direct discharge method. It is the simplest and most conservative approach which can be realized by the installation of a bare copper shielding wire along the pipeline. As a kind of grounding measure, it can decrease potentials with respect to the remote earth to the acceptable levels.

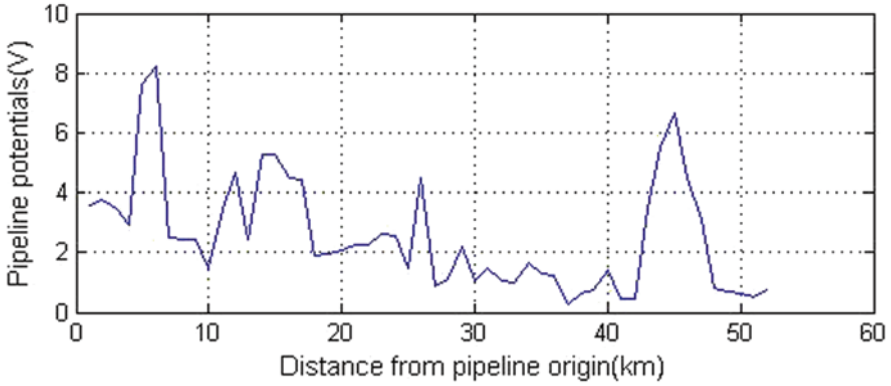


Fig. 48.5 Potentials along the gas pipeline with Mitigation Method 1

The specific plan is to install a 50-m bare copper shielding wire for every 110 m of the pipeline. The pipeline GPR is calculated and shown in Fig. 48.5. The potentials of the entire pipeline drop below 10 V.

The main cost of this method contains the purchase fee of bare copper wires and the construction cost of laying wires. The price of bare copper wires varies largely, and the common price is about RMB60/kg. Although this measure is more economical and effective for pipelines without cathode protection, it may fail to work for those with cathode protection, primarily because the cathode protection current discharges through the drainage network, which increases the output of cathode protection equipment and reduces the cathode protection range.

Mitigation Method 2: The cascade copper grounding electrode with decouplers. As the decoupler can block discharge (DC) current, the mitigation system can be used in pipelines with cathode protection. This method shares the same principle as Mitigation Method 1 but with more costs due to the extra decouplers. The calculation result of the pipeline GPR with the installation of this mitigation system is shown in Fig. 48.6.

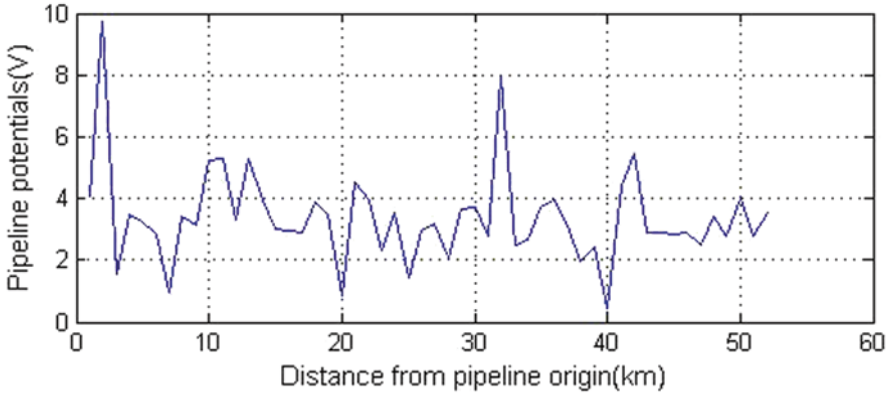


Fig. 48.6 Potentials along the gas pipeline with Mitigation Method 2

The specific plan is to install a 500-m bare copper wire at the points where the pipeline is very close to the transmission line and the pipeline potentials are extremely high, and the decoupler is between the pipeline and the copper wire. As the pipeline GPR is reduced to less than 10 V, it will not cause interference to the existing pipeline cathode protection system. This design is suitable for the discharge of AC and lightning current with long service life and high reliability.

Inducted AC voltages on gas pipeline are far beyond the criteria due to the present of the nearby 750-kV transmission line; as to the issue, the combinative mitigation measures are recommended.

48.4 Conclusion

- When the transmission line is parallel with the pipe along 10 km and the distance between them is above 473.33 m, the maximum pipeline potential will not exceed 10 V.
- Due to the long-distance parallel and small-angle crossover of the 750-kV transmission line and the gas pipeline, the potentials are quite high along the entire pipeline and the maximum value is above 100 V.
- The distance between the transmission line and the pipeline is suggested to be increased if possible. The shielding and discharging measures along the pipeline should be strengthened to meet the requirement that the pipeline GPR does not exceed 10 V during steady-state operation conditions.
- Small-angle crossover of the transmission line and pipeline should be avoided; if the crossover is inevitable, the vertical cross scheme is recommended.

References

1. Ma J, Dawalibi FP, Southey RD. Computation and measurement of electrical interference effects in aqueducts due to a nearby parallel transmission line, in Proc Int Symp Electromagn Compat, Beijing, IEEE; 1997. p. 409–12.
2. Lee S, Eun C, Cho PD. Analysis on the shielding effect by metallic pipe line against electromagnetic induced voltage, ICMMT Chengdu, IEEE; 2010. p. 1837–40.
3. Southey RD, Dawalibi FP, Vukonich W. Recent advances in the mitigation of AC voltages occurring in pipelines located close to electric transmission lines. IEEE Trans Power Deliv. 1994;9(2):1090–7.
4. Stet D, Czumbil L, Ancas L. Investigation of electromagnetic interferences issues, in Proc Int Symp ATEE, Bucharest, IEEE; 2013. p. 1–6.
5. Bortels L, Deconinck J, etc. A general applicable model for AC predictive and mitigation techniques for pipeline networks influenced by HV power lines. IEEE Trans Power Deliv. 2006;21(1):210–7.
6. SY/T0087.1–2006. The evaluation standard of steel pipe and storage tank corrosion-The direct assessment of buried steel pipeline external corrosion; 2006 (In Chinese).

Chapter 49
Online Monitoring System of the Grounding Status of Transmission Line Towers

Wei Li, Hailiang Lu, Zhiqiang He, Changcheng Zhu, Jiangtao Quan and Xishan Wen

Abstract The grounding status of transmission line towers is very important to the security and stability of the transmission line. Nevertheless, it is unfeasible and hard to carry out periodical manual checking of the grounding status in field conditions of the transmission line tower so as to find the grounding fault in time. Aiming at the problems, an online monitoring system in terms of the grounding status of transmission line towers is established. The system is composed of the measure device and the server, in which the measure device is based on the clamp meter method to send the tower grounding status to the server automatically and periodically. The measurement principle is simulated and analyzed in this chapter. Moreover, the prototype is tested in the laboratory environment, and the monitoring system is practically tested on the transmission towers. The result shows that the accuracy of the measure device is no less than 93.96% when grounding resistance is between 0.3 and 100Ω. Furthermore, the practical application shows that the system is effective and can be useful in the application of lightning protection.

Keywords Grounding · Tower grounding status · Clamp meter method · Online monitoring

49.1 Introduction

Lightning strikes to transmission lines are the primary cause of transmission-line outages [1]; moreover, the tower grounding state including the tower grounding resistance and the down-lead conducting state largely prevails as the main mechanism responsible for back flashover [2]. In this sense, periodical checking of the

H. Lu (✉) · W. Li · X. Wen
School of Electrical Engineering, Wuhan University, 430072 Wuhan, China
e-mail: hailiangLu@whu.edu.cn

Z. He
Hunan Electric Power Test & Research Institute, 410007 Changsha, China

C. Zhu · J. Quan
Electric Power Research Institute of Hubei Power Grid Corporation, 430077 Wuhan, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_49

grounding status of the transmission line tower is of great significance for lightning protection.

Unfortunately, it is a complex assignment to measure the tower grounding resistance. The traditional tower grounding resistance measurement is generally based on the three-electrode method by using the digital earth resistance measurement instrument. The tower down lead must be untied, and the measuring wire and electrodes should be arranged, which will bring about heavy workload and high environmental requirements [3]. Comparatively speaking, the measurement with the clamp meter is considered as a simpler approach. The measuring wire and the electrodes are not needed. Besides, the tower down lead should be untied with only one left, and the clamp meter should be clamped on it [4]; moreover, a portable grounding impedance meter based on a digital signal processor (DSP) is published [5]. The clamp method or the portable device may greatly reduce the workload of the measurement of the tower grounding resistance; however, it is still unfeasible to periodically measure the tower grounding resistance in field conditions by regular engineering personnel; and few studies have been carried out in terms of the tower grounding down lead solder fault and corrosion which can also cause the tower grounding.

With regard to the problems, a transmission line model strictly following the design drawing is established and simulated. On this basis, the whole monitoring system is designed and realized. Finally, the precision of the measure device is tested and compared with a high-precision digital multimeter in the laboratory environment; and the system is practically verified on the transmission towers.

49.2 Measurement Principle

The measurement principle derived from the clamp meter method is shown in Fig. 49.1.

The power coil and measure coil are set on the tower down lead. Energize the power coil to generate the induced voltage *e,* and induce the current *i* in the circuit. The loop resistance *z* is given by

$$Z = e / I. \tag{1}$$

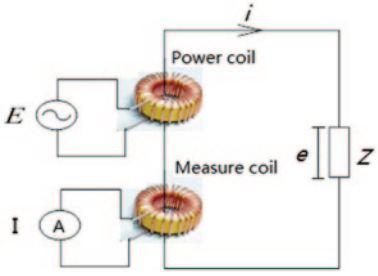


Fig. 49.1 Schematic diagram of measurement principle

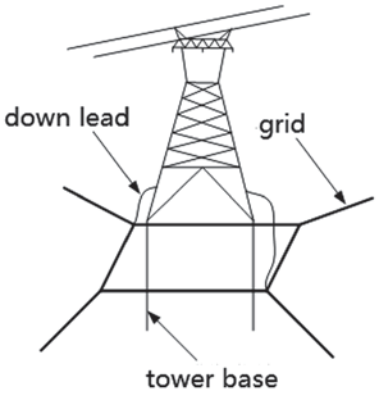


Fig. 49.2 Schematic diagram of tower grounding system

The induced voltage e is given by

$$e = E / n_1, \tag{2}$$

where n_1 are the power coil turns and E is the energizing voltage.

The induced current i is given by

$$i = n_2 I, \tag{3}$$

where n_2 are the measure coil turns and I is the current flow through the measure coil.

The loop resistance measurement result is given by Eqs. (1)–(3).

$$Z = e / i = \frac{E}{n_1 n_2 I} \tag{4}$$

As to a common transmission tower shown in Fig. 49.2, the definition of loop resistance Z depends on the tower down lead situation:

1. Both the down leads are normal: Z is the loop resistance of the tower body, the down leads, and the grounding grid circuit.

The loop resistance is very small in theory. But when the solder fault, the contact fault, and the corrosion exist, it could be much larger than normal.

2. The down lead with coils is normal and the other down lead is untied: According to the clamp meter method, Z is the loop resistance of the multi-towers which is approximate to the tower grounding resistance [3].
3. The down lead with coils is untied: The circuit is open and Z is infinite.

The simulation and actual test result is shown in Chap. 4.

49.3 Implemented Prototype

The prototype of the monitoring system must consider the field applications of the measurement devices, namely, long battery life, robustness, hardware protection, data transmission, storage, etc.

49.3.1 Hardware Setup

The monitoring system is composed of the measure devices and the server, in which each measure device is designed to be modular to improve the upgrade capabilities and easy maintenance. The system structure is shown in Fig. 49.3.

In consideration of the field condition, the measure device is powered by solar panel. And the electric storage is a 12-V lithium battery which charges under adequate light, and supplies the measure device. The charge controller is added in the system; the major functions of which are overcharge protection and over-discharge protection. The WRA1215ZP chips are used as DC-DC converters to ensure the sine-ware generator, filter and protection circuit, and control center work properly.

The sine-wave generator is the signal source for the measure system. An International Computer Limited (ICL) 8038 chip is used which puts out the sine-wave with a 3-V amplitude. The sine-wave flows through the power coil and generates induced voltage and current in the under-testing circuit. The power coil turns are 100 and the measure coil turns are 1500. The coil turn numbers are determined by experiment to get best accuracy. And the voltage and current measurement results are discretized by an analog of 12-bit, 100-ksps, and 8-channel-to-digital converter.

To prevent the lightning strike from threatening the measure device safety, the transient voltage suppressor has been used in the protection circuit.

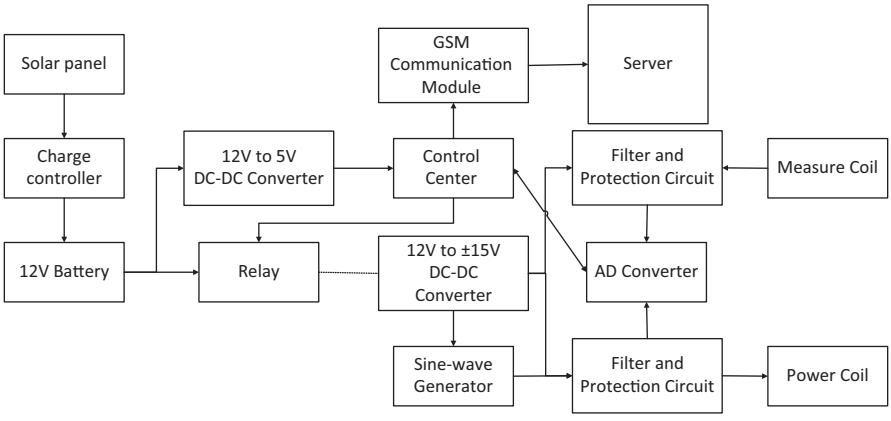


Fig. 49.3 Block diagram illustrating the system structure

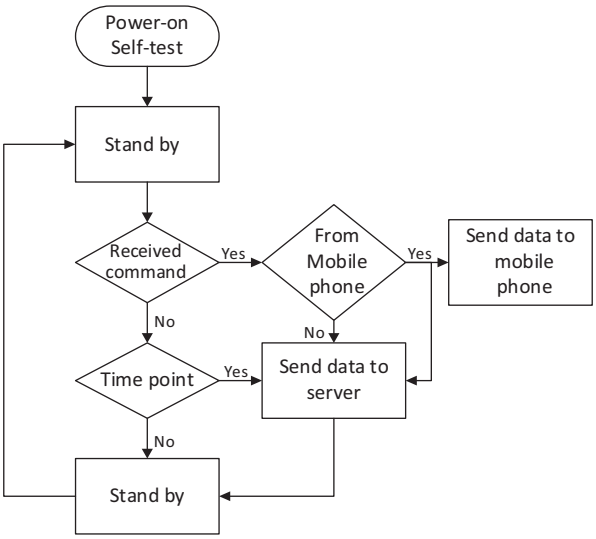


Fig. 49.4 Block diagram lllustrating the system procedure

The measurement devices are mostly working in field condition so that the communication between the measurement devices and the server should be wireless. In this prototype, the data transmission and the command sending are using the global system for mobile (GSM) communication system.

49.3.2 System Procedure

The devices automatically measure and send results to the server with a fixed cycle and standby during the rest of the time to save energy; but the command control is available at any time. The command can be sent to measure the server and any mobile phone. As to the field staff convenience, the measurement results are sent to both the server and the mobile phone when the command is sent by a mobile phone. The procedure of the system is shown in Fig. 49.4.

49.4 Simulation and Testing Results

49.4.1 Simulation Results

The tower model shown in Fig. 49.5 is built by strictly following the tower design sketch, including the conductor radius, length, materials, etc.

There are altogether 31 towers with a 500-m span in the whole transmission line.

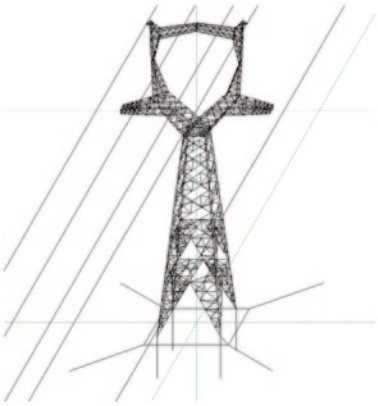


Fig. 49.5 Diagram of tower model

Energize 6-mV voltage in the circuit and measure the current. The earth is considered homogeneous with 500-Ω·m resistivity, the tower base concrete resistivity is 1000Ω·m, the solder fault and corrosion condition are approximated by adding a resistivity conduct of 0.5 m long and 0.01Ω·m to the circuit. The interference is to energize 606-A three-phase current in the transmission line; and the energizing frequency is 50 Hz.

Table 49.1 illustrates that the down lead condition has significant influence on the measurement result. When both of the down leads are normal, the measurement current circuit is down lead, grid, the other down lead, tower body; the resistance is at milli-ohm level and hardly has any current flow through the overhead grounding wire. The corrosion and solder fault raise loop resistance to ohm level with very little current flowing through the overhead grounding wire. When the down lead without coils is untied, the circuit becomes totally different and the result is the tower grounding resistance. Under this condition, the current flows through the overhead grounding wire and into ground from other towers. Finally, the interference could be even more than the measurement current under condition I, which

Table 49.1 Simulation results of loop resistance measurement

Voltage level (kV)	Condition	Current in down lead (mA)	Current in overhead grounding wire (one side)	Loop resistance (Ω)
500	N	71	0	85m
220	N	91	0	66m
500	I	110	11.6A	none
220	I	180	11.6A	none
500	F	3.4	0.01mA	1.76
220	F	3.2	0.01mA	1.88
500	U	0.83	0.41mA	7.26
220	U	0.77	0.39mA	7.77

N down lead condition is normal without interference, *I* no down lead energized with only interference, *F* down lead has corrosion and solder fault, *U* the down lead without coils untied

Fig. 49.6 View of the prototype

may cause a significant error; therefore, the device measurement frequency should be different from 50 Hz, and a filter is necessitated.

49.4.2 Practical Test

The measurements have been carried out to test the performance of the prototype as shown in Fig. 49.6.

49.4.2.1 Laboratory Test

The first tests were carried out in the laboratory to measure different loop resistances. The results were compared with the measurement results by the high-precision digital multimeter, as shown in Table 49.2.

Table 49.1 illustrates that the maximum relative error between the prototype and the high-precision digital multimeter is 6.04%. The accuracy satisfies the requirement.

Table 49.2 Comparison of measurement results

Multimeter (Ω)	Prototype (Ω)	Relative error (%)
0.314	0.321	2.16
0.522	0.518	0.67
0.863	0.875	1.39
1.042	1.105	6.04
2.143	2.118	1.16
5.114	5.187	1.42
10.140	10.382	2.39
20.250	20.324	0.37
50.560	51.032	0.93
100.100	99.873	0.23

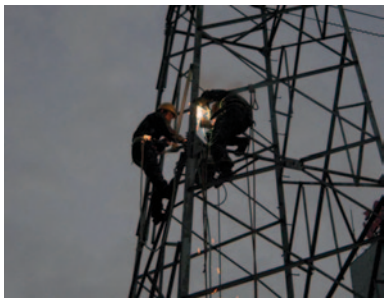


Fig. 49.7 View of prototype installation

Table 49.3 Practical measurement result

Tower	Condition	Loop resistance
1	N	0.111Ω
1	U	8.234Ω
1	D	OR
2	N	0.460Ω
2	U	3.819Ω
2	D	OR
3	N	OR

N down lead condition is normal, *U* the down lead without coils untied, *D* both down leads disconnected, *OR* out of range

49.4.2.2 Practical Application

The prototype was tested on several transmission towers in China. (Fig. 49.7)

The measurement results are shown in Table 49.3.

Table 49.3 illustrates that the loop resistance at normal condition is always larger than ideal case because of the contact resistance; however, it is still much smaller than other conditions. The measurement result of tower 3 indicates the existence of grounding fault. The excavation at tower 3 confirms the down lead disconnection as shown in Fig. 49.8.

Fig. 49.8 View of grounding fault of tower 3

49.4.3 Discussion

The system is proved to be effective and convenient in the practical tests. The field staff can get tower down lead status at anytime and anywhere by sending a command message, but the field staff must untie other down leads of the concerned tower to get data in terms of the tower grounding resistance because the system cannot get the down lead status and the tower grounding resistance at the same time, which should be solved in further studies.

49.5 Conclusion

This chapter analyzes the measurement of loop resistance based on the clamp meter method and simulates the loop resistance at different conditions. On this basis, a new transmission tower grounding status monitoring system is designed and implemented. The devices are designed to be modular in consideration of the field application as well. Preliminary tests have been carried out in the laboratory with the results to be compared with the measurement results by a high-precision digital multimeter in order to verify their accuracy; moreover, the practical tests on transmission towers demonstrate that the device is effective and useful in the application of lightning protection. Furthermore, the modular design of the device is conducive to improve the upgrade capabilities and easy maintenance.

References

1. Furukawa S, Usuda O, Isozaki T, et al. Development and application of lightning arresters for transmission lines. IEEE Trans Power Deliv. 1989;4(4):2121–9.
2. Visacro S. Direct strokes to transmission lines: considerations on the mechanism of overvoltage formation and their influence on the lightning performance of lines. J Lightning Res. 2007;1(1):60–8.
3. Lin YQ. Discussion on grounding resistance measurement of transmission tower. High Volt Eng. 2004;30(8):60–1. (In Chinese).
4. Chen JZ, Shi XG. The study of testing methods for tower grounding resistance. Power Syst Technol. 2006;30(82):495–8. (In Chinese).
5. Rodrigues BD, Visacro S. Portable grounding impedance meter based on DSP. IEEE Trans Instrum Measure. 2014;63(8):1916–25.

Chapter 50
Application of the Leader Progression Model in the Insulation Flashover Criterion for Lightning Performance Calculation

Xiaoyue Chen, Xishan Wen and Lei Lan

Abstract It is rather important to exactly evaluate the lightning withstanding level of the transmission line for the system service and the stable operation. In the course, the intersection method is usually adopted as the insulation flashover criterion in the practical project. But due to the difference of the nature lighting waveform and the standard lighting waveform (1.2/50 us), and the effect of the reflection and refraction in the lightning progression process, it may cause a large error by using the intersection method based on the volt-second characteristic curve, which is obtained from the standard lighting impulse in engineering. Several methods of estimating the insulation flashover are in use today. In general, these methods can be divided into those based on the conclusion of insulation breakdown and those derived from the breakdown process. The leader progression model used by CIGRE is one of those methods. In this chapter, several estimation methods of the insulation flashover are investigated and compared. A new insulation flashover criterion is proposed herein on the basis of the experimental results of the leader progression model and the CIGRE insulation flashover criterion. The application of this simulation model and the flashover time under standard lightning impulses are calculated to obtain the volt-second characteristic curve, which corresponds to the results obtained by the experiment.

Keywords Leading progression model · Flashover criterion · Lightning impulse · Volt-second characteristic

50.1 Introduction

The safety of transmission lines directly affects the safety and reliability of the power system. Thus it will be of great significance to calculate the lightning overvoltage exactly in the case of lightning strike on the transmission lines because it is one of the main faults in the power system.

X. Chen (✉) · X. Wen · L. Lan
School of Electrical Engineering, Wuhan University, 430000 Wuhan, China
e-mail: ringgle1987@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_50

The lightning protection calculation of the transmission line is usually based on the conclusion of insulation breakdown to determine whether the insulation flashover happens by comparing the waveform or the amplitude of overvoltage at the end of the insulator string. The intersection method is a representative one among these kinds of methods by comparing the insulator volt-second characteristic curve under the standard lightning surge 1.2/50 (us) with the waveform of overvoltage at the end of the insulator string [1, 2]. Once the waveform of overvoltage and the volt-second characteristic curve of the insulator intersect, the insulator string will be determined to flashover. The intersection method is widely used in the lightning protection calculation of the transmission line because of its clear physical concept and convenience of application to the engineering calculation.

The measurement and simulation indicate that lightning overvoltage at the end of the insulator string is actually far from a standard wave, but has a short tail [3, 4]. The discharge voltage of the insulator string under short-wave impulse voltage will be higher than that under the standard wave, and the short tail wave will make the volt-second characteristic curve upward; therefore, when the intersection method is adopted in lightning protection calculation of the transmission line, the lightning withstanding level will be underestimated and the lightning intruding overvoltage will be underestimated when it is adopted in the lightning protection calculation of substation.

This chapter adopts the leader progression method to determine the flashover process of the insulator string. First, the flashover model of the insulator string is established; another program is developed based on this model to calculate the volt-second characteristic curve of the insulator string; then the effects of the average field intensity of the steam and leader channel are discussed; finally, the volt-second characteristic curve obtained by simulation and experiment is compared to verify that the application of the leader progression method to the insulator string flashover criterion is effective.

50.2 Criterion of Insulation Flashover based on the Leader Progression Method

The criterion of insulation flashover based on the leader progression method considers the physical process of air gap discharge and combines the insulation flashover process under impulse voltage with the air gap flashover process of the same length. As the leader progression method of air gap discharge consists of several successive stages, including corona inception, streamer progression, leader progression, and main discharge, the insulation breakdown time is made up of several components [5]:

$$T_b = T_p + T_s + T_i + T_l + T_g, \tag{50.1}$$

where

T_p is the corona inception time,
T_s is the streamer progression time,
T_i is the ionic birefringence wave progression time,
T_l is the leader progression time, and
T_g is the gas heating time.

The ionic birefringence wave progression time T_i and the gas heating time T_s are very short when compared to other time values; so T_i and T_g can be ignored in the calculation. The corona inception voltage is far lower than the flashover voltage in a highly nonuniform electric field; so, the corona inception time T_p is negligible too. The insulation breakdown time can be calculated by the sum of T_s and T_l time values.

Streamer progression time can be calculated as below [6]:

$$T_s = \frac{1}{k_1 \dfrac{E}{E_{50}} - k_2}, \tag{50.2}$$

where E is the maximum electric field intensity before insulation breakdown, kV/m; E_{50} is the electric field intensity at the end of insulation under the 50% flashover voltage, kV/m; k_1 and k_2 are the factors of streamer progression time, which are recommended to be 1.25 and 0.95, respectively, as obtained from the air gap discharge experiment by Pigini.

The leader progression time can be calculated from its speed which is recommended by CIGRE as below [6],

$$\frac{dx}{dt} = ku(t)\left[\frac{u(t) - E_z x}{D - x} - E_{l0}\right], \tag{50.3}$$

where x is the length of leader, m; $u(t)$ is the voltage at the end of insulation, kV; D is the length of insulation, m; E_{l0} is the threshold electric field intensity of leader progression, kV/m; E_z is the average electric field intensity of the leader channel; k is the factor of leader progression speed. E_s and k are associated with the form of the insulator and the polarity of impulse voltage. According to CIGRE WG33.01 results, typical values obtained from experiments are listed in Table 50.1.

The leader progression speed will increase rapidly before the insulation breakdown, and the head of the leader may arrive at the other end instantly. The length of the final jump is recommended by Rizk as below [7]:

$$h_f = \frac{3.89}{1 + \dfrac{3.89}{D}}, \tag{50.4}$$

where D is the length of insulation.

50.3 Procedure of Insulation Flashover

The calculation of the insulation breakdown time is accomplished by T_s and T_l. Then the insulation breakdown process can be reproduced by the simulation program based on the leader progression model. The flowchart of the criterion of insulation flashover is shown in Fig. 50.1.

The voltage at the end of the insulation can be obtained by an electromagnetic transient program. The streamer develops once the electric field intensity exceeds a threshold value, which is 350 kV/m. And T_s can be calculated by Eq. 50.2 by the maximum electric field intensity before the insulation flashover. Then, the voltage amplitude keeps increasing, and the electric field intensity exceeds the leader's initial electric field intensity E_{l0}, which is listed in Table 50.1, and it is enough to let leader develop continuously. The leader progression speed can be calculated by Eq. 50.3. The insulation will finally flash over once the length of the remaining gap is less than that of the final jump.

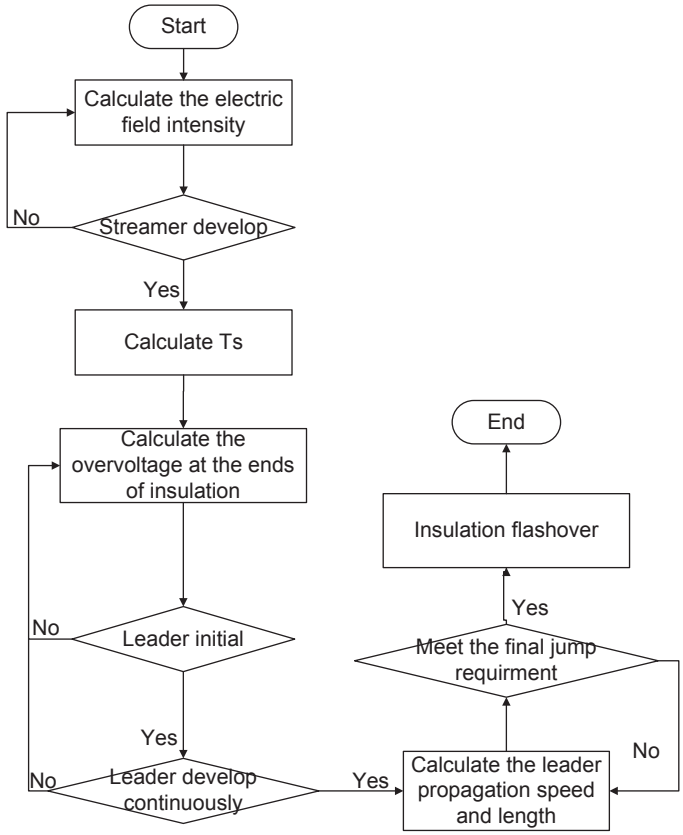


Fig. 50.1 Flowchart of the criterion of insulation flashover

Table 50.1 Threshold electric field intensity of leader progression and the factor of leader progression speed

The form of insulator	The polarity of impulse voltage	The threshold electric field intensity of leader progression E_{l0}	The factor of leader progression speed k
Suspension insulator	Positive	520	1.2
Line post insulator	Negative	670	0.8

50.4 Parameter Discussion

It is clear that the head of the leader is extremely active in the ionization area, and the head of the leader actually performs the streamer characteristics. If the leader progresses forward, the head of the leader with streamer characteristics must progress preferentially; hence, the threshold electric field intensity of leader progression E_{l0} in Eq. 50.3 can replace the average electric field intensity of streamer E_s (Table 50.2).

Les Renarieres laboratory analyzed and summarized E_s and E_z from the experiment data of the 5, 10, and 13.5 m air gap breakdown process as shown in Table 50.2 [8].

The criterion of insulation flashover based on the leader progression method is adopted to simulate the experiment of the volt-second characteristic curve. The length of insulation and the voltage waveform are kept constant while the amplitude of voltage is increased to make an insulation flashover. Record the relationship between the voltage peaks that appeared in the flashover process and the breakdown time to obtain the volt-second characteristic curve.

Darveniza put forward an empirical formula to describe volt-second characteristics of the insulator string by employing constant length as follows [9]:

$$U_{s-t} = 400L + 710L / t^{0.75}, \tag{50.5}$$

where L is the length of the insulator string, m; t is the time, us.

The effect of different E_s and E_z has been considered on a 10-m air gap. The curve obtained from simulation and the experiment are compared as shown in Figs. 50.2 and 50.3.

The results show that as the average field intensity of the leader channel increases, when the voltage on the residual air gap is reduced, the flashover voltage of the insulator string is increased, and the volt-second characteristic curve has moved

Table 50.2 Threshold electric field intensity and average electric field intensity of the leader channel

D/m	5	10	13.5
E_s/kV/m	520	510	480
E_z/kV/m	160	100	80

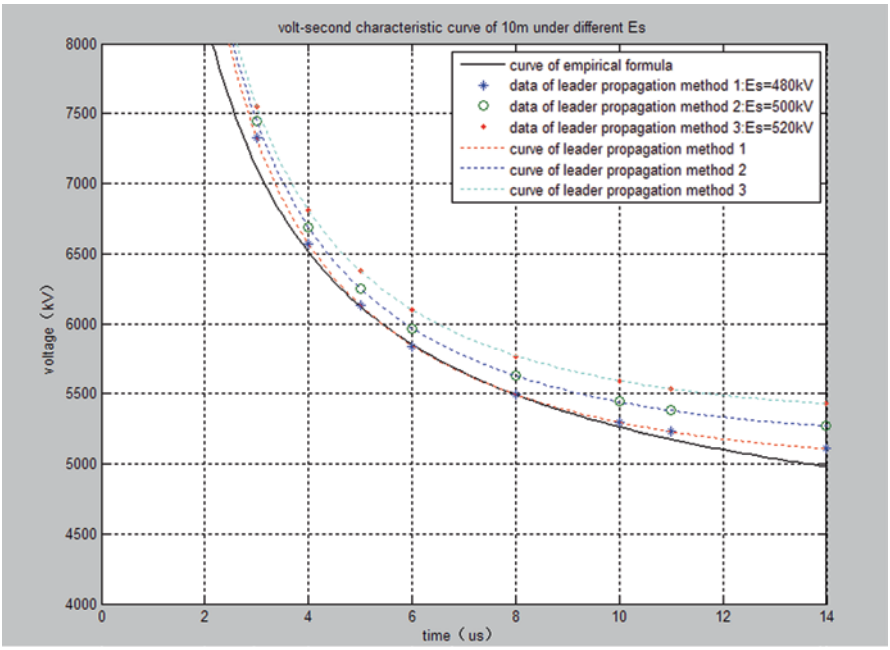


Fig. 50.2 Volt-second characteristic curve of 10 m under different E_s values

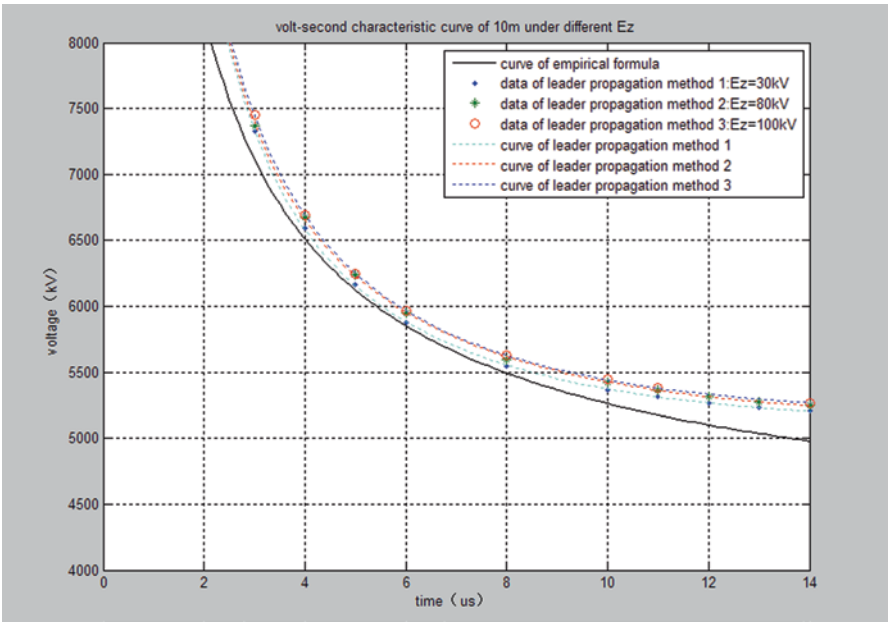


Fig. 50.3 Volt-second characteristic curve of 10 m under different E_z values

up. As the average field intensity of the streamer increases, when the streamer and leader initial threshold voltage are increased, the flashover voltage of the insulator string is increased, and the volt-second characteristic curve has moved up too.

50.5 Comparison of Simulation and Experiment Results

Wuhan High Voltage Research Institute has gained the volt-second characteristic of 31 pieces of XWP-16 insulator strings from the experiment [10]. The structure height of the insulator string is 4.805 m. The flashover process is reproduced by simulation, and the volt-second characteristic curve obtained from simulation is compared with that from the experiment.

The values of E_s and E_z are taken, respectively, to be 520 and 160 kV/m. When the applied voltage is 4580 kV, the insulator string flashes over at 2 us, and the maximum speed of leader progression is 5.543 m/us. In fact, the flashover of insulation is the surface flashover, whose flashover voltage is lower than the flashover in the air, and even they have the same flashover distance; in this sense, the slight difference in the results between simulation and the experiment in Fig. 50.4 is caused by lower flashover voltage of surface flashover; besides, the simulation of volt-second characteristic curves is slightly higher than that of the experiment. And they are generally consistent.

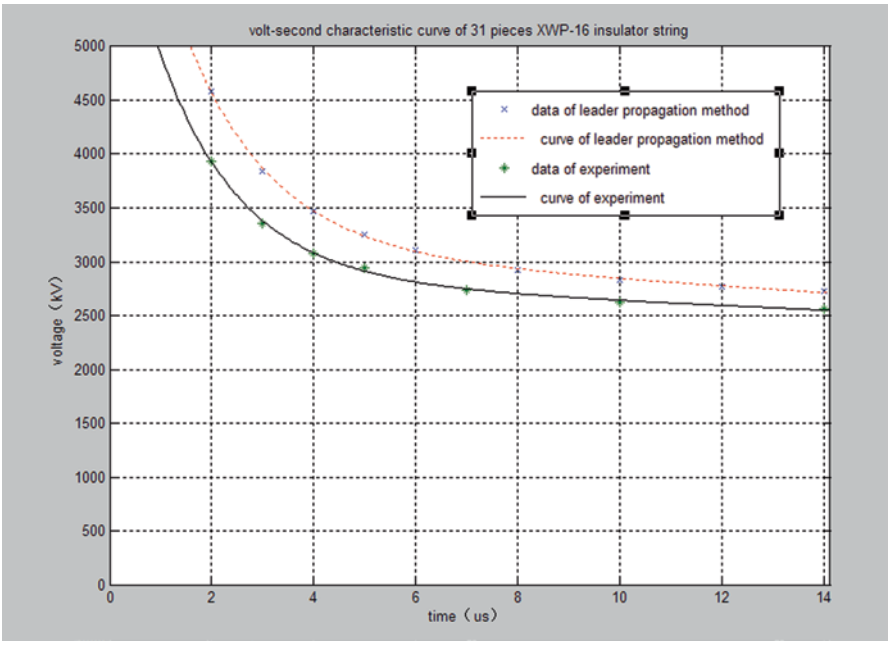


Fig. 50.4 Volt-second characteristic curve of a 31-piece XWP-16 insulator string

50.6 Conclusion

In this research, a numerical calculation program of the insulation flashover criterion based on the leader progression method is applied to reproduce the process of insulation flashover. With this criterion, the effects of the average field intensity of the steam and leader channel are discussed by comparing the volt-second characteristic curve obtained from simulation and experiment, respectively. With the rational parameter value, the lightning flashover performance of 31 pieces of XWP-16 insulator strings for 500 kV transmission lines is calculated here with the calculation results compared with experimental results. This indicates that the criterion is reasonable.

References

1. IEEE Working Group on Estimating Lightning Performance of Transmission Line. A simplified methods on estimating lightning performance of transmission line. IEEE Trans Power Appar Syst. 1985;104(4):919–31.
2. Hui-sheng YE, Jun-jia HE, Hua LI. Simulation of overvoltage and flashover caused by lightning stroke at towers of HVDC transmission line. Power Syst Technol. 2005;29(21):31–5. (In Chinese).
3. Carrus A, Clnieri E, Fumi A, et al. Short tail lightning impulse behaviour of medium voltage line insulation. IEEE Trans Power Deliv. 1999;14(1):218–27.
4. Ancajima A, Carrus A, Clinier E, et al. Breakdown characteristics of air spark-gaps stressed by standard and short-tail lightning impulses: experimental results and comparison with time to sparkover models. J Electrost. 2007;65(5/6):282–8.
5. Shindo T, Suzuki T. A new calculation method of breakdown voltage time characteristics of long air gaps. IEEE Trans Power Apparatus Syst. 1985;104(6):1556–63.
6. Pigini A, Rizzi G, Garbagnati E, et al. Performance of large air gaps under lightning overvoltage: experimental study and analysis of accuracy of predetermination methods. IEEE Trans Power Deliv. 1989;4(2):1379–87.
7. Rizk FAM. A model for switching impulse inception and breakdown of long air-gaps. IEEE Trans Power Deliv. 1989;4(1):596–606.
8. Les Renardieres Group. Positive discharges in long air gaps at Les Renardieres 1975 results and conclusions. Electra. 1977;53:31–152.
9. Darveniza M, UHV Transmission Lines and Lightning. Brisbane: University of Queensland Brisbane (1977).
10. Hui Hu. Research on Lightning Protection of 500 kV Double-circuit Transmission Lines. Beijing: Tsinghua University (1997) (In Chinese).

Chapter 51
The Risk Index System of Power Transformer Life Cycle Cost and Its Weight Determination

Yuqin Xu, Zheng Ren, Xiangling Zhan, Tong Li, Weitao Hu, Guohua Qiao and Qing Xie

Abstract In order to evaluate the risk of transformer life cycle cost, we need a set of key indicators to reflect the cost risk of the transformer life cycle; thus, it is important to set up a scientific risk index system for the purpose of risk assessment. This chapter establishes a more comprehensive cost model in terms of the transformer life cycle and quantifies the uncertain factors of power transformer life cycle cost management; then we can obtain a risk assessment index system of the transformer life cycle cost. We apply an analytic hierarchy process to the cost risk indicators in respect of the transformer life cycle so as to construct the judgment matrix, get the weight coefficients of various risk indicators, and lay the foundation for risk assessment of transformer life cycle cost.

Keywords Power transformer · Life cycle cost · Index system · Analytic hierarchy process

51.1 Introduction

The power transformer is one of the most important power transmission and transformation equipment, which is in the hub status in the power system. It is the foundation for applying appropriate life cycle cost (LCC) management to the transformer for the safe and economic operation of the power system. With the introduction of the LCC theory to our country, extensive research studies have been carried out with remarkable results obtained. In the traditional investment decision of the transformer, electric power enterprises lack the understanding of a full life cycle which has not fully considered the total cost in the process of a full

Z. Ren (✉) · Y. Xu · X. Zhan · T. Li · Q. Xie
College of Electrical and Electronic Engineering, North China Electric Power University, 071000 Baoding, China
e-mail: renzheng6536@126.comw

W. Hu · G. Qiao
State Grid Hebei Electric Power Company, 050000 Shijiazhuang, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_51

life cycle, and there are uncertain factors affecting the whole LCC; therefore, the establishment of a risk index system of transformer LCC is of great significance in improving the reliability of power transformer operation and reliability of the power grid [1–3].

As to the modeling research on transformer LCC, scholars have made some research results at home and abroad. Tao Li made the cost decomposition in the whole substation equipment life cycle and established the substation equipment LCC model [4]. The LCC model is divided into five parts: investment cost, operation cost, maintenance cost, failure cost, and disposal cost [5]. Wies and Nilsson made the LCC analysis of equipment, which used different methods of simulation modeling [6, 7]. Xiaochu Luo proposed a variety of maintenance strategies and obtained their effect on LCC [8].

On the basis of the above research results, this chapter establishes a more comprehensive transformer LCC model and quantifies the uncertain factors of power transformer LCC management; then we can obtain a transformer LCC risk assessment index system. We apply analytic hierarchy process to transformer LCC risk indicators, construct the judgment matrix, get the weight coefficients of various risk indicators, and lay the foundation for risk assessment of transformer LCC.

51.2 Life Cycle Cost Model of Power Transformer

In this chapter, the transformer LCC is divided into seven parts: investment cost, operation cost, environmental cost, preventive test cost, maintenance cost, failure cost, and disposal cost. The LCC model equation is shown as follows:

$$LCC = C_{\mathrm{CI}} + C_{\mathrm{CO}} + C_{\mathrm{CE}} + C_{\mathrm{CS}} + C_{\mathrm{CM}} + C_{\mathrm{CF}} + C_{\mathrm{CD}}. \tag{51.1}$$

Among them, C_{CI} is the investment cost; C_{CO} is the operation cost; C_{CE} is the environmental cost; C_{CS} is the preventive test cost; C_{CM} is the maintenance cost. C_{CF} is the failure cost and C_{CD} is the disposal cost.

51.2.1 Investment Cost

The investment cost of the transformer is the cost of the equipment before operation; in this article, the investment cost is divided into planning design (C_{gh}), procurement (C_{cg}), construction (C_{az}), and other expenses (C_{qt}).

$$C_{\mathrm{CI}} = C_{gh} + C_{cg} + C_{az} + C_{qt}. \tag{51.2}$$

51.2.2 Operation Cost

The operation cost is the running cost of the transformer within the LCC, including the artificial maintenance cost (C_{wh}) and energy consumption (C_{nh}).

$$C_{\mathrm{CO}} = C_{nh} + C_{wh}. \tag{51.3}$$

51.2.3 Environmental Cost

With the rapid development of human society economy, the increasing environmental problems cannot be ignored. Global warming is a serious threat to the sustainable development of human society, and the gases which produce greenhouse are mainly CO_2, CH_4, NO_X, etc. At the same time, the electromagnetic radiation and the noise pollution of power transformer also bring some harm to the surroundings. This chapter divides the environmental cost into the greenhouse gases penalty cost, the electromagnetic radiation penalty cost, and the noise pollution penalty cost.

$$C_{CO2} = \sum_{t=1}^{N} (E_{LR} - E_{L0}) C_E \left(\frac{1+r}{1+R} \right)^{t-1} \tag{51.4}$$

$$C_{EM} = \sum_{t=1}^{N} \sum_{i=1}^{3} (E_{EMi} - E_{EM\,0i}) C_{Ei} \left(\frac{1+r}{1+R} \right)^{t-1}. \tag{51.5}$$

Among them, E_{LR} represents the annual actual cost, E_{L0} represents the annual normal cost, E_{EMi} represents the actual radiation, $E_{EM\,0i}$ represents the standard radiation limit, C_{Ei} represents the radiation exceeding penalty cost, r represents the rate of inflation, and R represents the social discount rate.

Noise pollution mainly includes the transformer construction noise and the electromagnetic noise of the run time. The noise pollution common valuation methods include protection fees, willingness to pay, and direct or indirect cost of the damaged method.

51.2.4 Preventive Test Cost

Preventive test cost includes the cost of preventing equipment from failure and regular preventive test cost with the equation shown as below:

$$C_{CS} = \sum_{i=1}^{\lfloor N/\Delta t_S \rfloor} C_{SA}(i\Delta t_S) \left(\frac{1+r}{1+R} \right)^{i\Delta t_S}. \tag{51.6}$$

Among them, $C_{SA}(i\Delta t_S)$ represents the preventive test year cycle.

51.2.5 Maintenance Cost

The maintenance cost refers to the transformer maintenance costs caused by the failure, including the overhaul cost and minor repair cost.

$$C_{CM} = \sum_{t=1}^{N} \lambda(t)[C_{MAO}(t)\rho(t) + C_{MAMR}(t)(1-\rho(t))]\left(\frac{1+r}{1+R}\right)^{t-1}. \tag{51.7}$$

Among them, $\lambda(t)$ represents the failure rate, $\rho(t)$ represents the probability of overhaul, $C_{MAO}(t)$ represents the preventive the overhaul cost, and $C_{MAMR}(t)$ represents the minor repair cost.

51.2.6 Failure Cost

This chapter is based on years of actual power outage time to measure the fault outage punishment cost, building a more comprehensive fault outage loss cost, mainly including three parts: failure to supply loss, repair fault cost, and penalty cost.

$$C_{CF} = \sum_{t=1}^{N} [aWT + \lambda(t)RC \bullet MTTR + (T - T_0)C_d]\left(\frac{1+r}{1+R}\right)^{t-1}. \tag{51.8}$$

Among them, a represents the electrovalence, W represents the annual interruption of power, T represents the annual actual power outage time, $\lambda(t)$ represents the failure rate, RC represents the average annual repair costs, $MTTR$ represents the average annual repair time, T_0 represents the annual plan outage time, and C_d represents the power outage time compensation fee.

51.2.7 Disposal Cost

Transformer disposal cost includes the scrap cost and salvage value.

$$C_{CD} = (C_R + C_V - C_{NS})\left(\frac{1+r}{1+R}\right)^{T-1}. \tag{51.9}$$

Among them, C_R represents the clearing expense, C_V represents the value loss expense, C_{NS} represents the expense of the recycling of the residual value.

51.3 Risk Index System of Power Transformer Life Cycle Cost

The principles of the risk assessment indicator system include the life cycle principle, the clear and practical principle, and the quantitative qualitative analysis principle.

According to the above principles of establishing an index system, set up a transformer LCC risk assessment index system. The LCC risk assessment index system consists of 7 first-level indexes and 38 secondary indexes. The index system is shown in column1, column 2, and column 4 in Table 51.1.

51.4 Application of the Analytic Hierarchy Process to Determine Weight

By 1–9 scaling to obtain judgment matrixes, we can apply the AHP method to get the LCC risk evaluation system index weights, as shown in column 3 and column 5 in Table 51.1.

The analytic hierarchy process is generally divided into three steps, taking the preventive test cost secondary risk factors in Table 51.1, for example, the process of calculating the weights is as follows:

1. Construct judgment matrix based on the expert assignment. For example, the judgment matrix of preventive test cost risk can be written as: $B_4 = \begin{pmatrix} 1 & 3 & 5 \\ 1/3 & 1 & 3 \\ 1/5 & 1/3 & 1 \end{pmatrix}$.
2. Verify the consistency of the judgment matrix B_4. After calculation, we get to know that this matrix meets the requirements.
3. Calculate the weight of B_4, and we obtain weight $w = (0.6370, 0.2583, 0.1047)$. It shows the weights of the test personnel operation risk, material price change risk, and technology risk. After a total sort, we can get the weight of the test personnel operation risk, the material price change risk, and the technology risk of $w = (0.0276, 0.0112, 0.0045)$.

Other primary and secondary index weights can also be obtained according to the above process. The result is shown in Table 51.1.

From Table 51.1, we can obtain the weights of 7 first-level indexes and 38 secondary indexes. According to the first-level index weights, we get to know

Table 51.1 Risk assessment index system of LCC

Column1	Column2	Column3	Column4	Column5
Evaluation target	First level indexes	First level indexes weights	Secondary indexes	Secondary indexes weights
LCC cost risk	Investment cost risk	0.1427	Security risk of planning and design	0.0114
			Equipment price floating risk	0.0056
			Technology risk	0.0140
			Investment risk	0.0491
			Risk of load forecasting precision	0.0232
			Equipment location risk	0.0317
			Personnel quality risk	0.0030
			Construction quality guarantee risk	0.0046
	Operation cost risk	0.0760	Operational risk	0.0308
			Load control risk	0.0143
			Legal risks	0.0025
			Management risk	0.0181
			Natural disaster risk	0.0067
			Energy and emissions reduction policy risk	0.0035
	Environmental cost risk	0.0509	Environmental pollution compensation risk	0.0288
			Technology risk	0.0028
			Policy risk of variability	0.0133
			Laws and regulations permit risk	0.0060
	Preventive test cost risk	0.0433	Test personnel operation risk	0.0276
			Material price change risk	0.0112
			Technology risk	0.0045
	Maintenance cost risk	0.2181	The uncertainty risk of repair	0.1114
			Repair effect risk	0.0137
			Personnel operating risk	0.0589
			Overhaul management risk	0.0254
			Technology risk	0.0087
	Failure cost risk	0.4440	The uncertain failure frequency risk	0.1819
			Loss of load risk	0.0412
			Sudden blackout risk	0.0168
			Emergency workers quality risk	0.1166
			Policy risk of variability	0.0255
			Power compensation risk	0.0620
	Disposal cost risk	0.0249	Equipment damage value risk	0.0046
			Price risk	0.0024
			Scrap application risk	0.0012
			Scrap approval risk	0.0007
			Idle supplies statistical risk	0.0061
			Idle material disposal risk	0.0100

that the failure cost risk is the highest; at the same time, both the investment cost risk and the maintenance cost risk are very high. In the same way, we can obtain the uncertainty risk of repair; and the uncertain failure frequency risk and emergency workers' quality risk are very high according to the 38 risk factors in Table 51.1.

The weights mainly reflect the effect of various risk indexes on the transformer LCC, which lays the foundation for risk assessment of transformer LCC.

51.5 Conclusion

The chapter establishes a set of cost risk assessment systems in terms of the transformer life cycle, including transformer investment cost risk, operation cost risk, environmental cost risk, preventive test cost risk, maintenance cost risk, failure cost risk, and disposal cost risk. The system is scientific and reasonable. We apply the analytic hierarchy process to the risk indicators of transformer LCC, construct the judgment matrix, get the weight coefficient of 7 first-level indexes and 38 secondary indexes, and lay the foundation for the risk assessment of transformer LCC.

Acknowledgments This work was supported by the National Natural Science Foundation of China (51307060), Colleges and Universities in Hebei Province Science and Technology Research Project (Z2013086), Fundamental Research Funds for the Central Universities (2014MS84), State Key Laboratory of Alternate Electrical Power System with Renewable Energy Source (LAPS14011), and the State Grid Corporation Project (kj2013–065).

References

1 Liao R, Yang L, Zheng H, et al. Reviews on oil-paper insulation thermal aging in power transformers. Trans China Electrotech Soc. 2012;27(5):1–10 (In Chinese).
2 Yin J, Zhu Y, Yu G, et al. Fault diagnosis of transformers based on Gaussian process classifier. Trans China Electrotech Soc. 2013;28(1):158–64 (In Chinese).
3 Liu Y, Ma L, Wu L, et al. Economic life model of power transformer and its application. Power Syst Technol. 2012;36(10):235–40 (In Chinese).
4 Li T, Ma W, Huang X. Power transformation equipment management based on life cycle cost theory. Power Syst Technol. 2008;32(11):50–3 (In Chinese).
5 Bojic M, Dragicevic S. Optimization of steam boiler design. Proc Inst Mech Eng, Part A: J Power Energy. 2006;220(6):629–34.
6 Wies RW, Johnson R, Agrawal AN. Simulink model for economic analysis and environmental impacts of a PV with diesel-battery system for remote village. IEEE Trans Power Syst. 2005;20(2):692–700.
7 Nilsson J, Bertling L. Maintenance management of wind power systems using condition monitoring systems-life cycle cost analysis for two case studies. IEEE Trans Energy Convers. 2007;22(1):223–9.
8 Luo X, Li L, Wei Z, et al. Applications of life cycle cost theory in decision-making of investment for distribution transformers renovation. Power Syst Technol. 2011;35(2):207–11 (In Chinese).

Chapter 52
Application of Differentiated Lightning Protection to Typical Transmission Lines in Mountainous Areas

Yabing Zhou, Jianzhi Lin, Luo Ge, Qinhua Feng and Guobiao Ning

Abstract The lightning protection transformation is strongly needed because the transmission lines in mountainous areas have been frequently subject to the lightning damage. This chapter carries out research studies on the lightning protection of transmission lines in mountainous areas by the technology of differentiated lightning protection; at the same time, the lightning risk assessment using differentiated lightning protection is conducted as to 110 kV transmission line with the most frequent flashovers. The factors of terrains and topographic features of towers, lightning activities, and insulation configuration have been considered in the course of risk assessment. The scheme of lightning protection transformation is thus proposed based on the assessment results technically and economically.

Keywords Differentiated lightning protection · Transmission lines · Lightning protection transformation · Trip-out rate · Line arrester

52.1 Introduction

The lightning strike is the primary cause of the transmission line fault. The line's trip-out of State Grid Corporation increases every year from 2005 to 2010 with 40 % of such trip-out caused by lightning, especially the transmission lines in mountainous areas [1]. They are mostly in the mountains of big ridge and complex terrain. Experience has shown that the lightning trip rate of transmission lines in southern mountainous area is higher than that of the electricity supply with relevant safety and reliability as well as the power grid of safe operation [2].

In recent years, China has carried out a lot of research studies to improve the situation of lightning protection in mountainous areas [3, 4]. The technology of dif-

L. Ge (✉)
School of Electrical Engineering, Wuhan University, 430072 Wuhan, China
e-mail: 568574828@qq.com

Y. Zhou · J. Lin · Q. Feng · G. Ning
Qingyuan Power Supply Company, 511500 Qingyuan, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_52

ferentiated lightning protection was put forward which considered the differences of various factors such as distribution of lightning activity, topography, tower structure, lightning accident, etc. Then, it recommended appropriate lightning protection transformation program [5, 6]. As a new method, whether it can achieve a corresponding effect or not remains to be demonstrated in practice.

In this chapter, it takes the lighting protection transformation in some mountainous areas as an example. It researches the lightning protection transformation of transmission lines with the method of differentiated lightning protection. Finally, it provides a cost-effective rehabilitation program [7]. The work provides a strong argument for the application of the technology in engineering practice.

52.2 Technology of Differentiated Lightning Protection

It is necessary for the transmission lines in mountain areas with serious lightning to transform the lightning protection which is a very complex process. In this chapter, the technology of differentiated lightning protection is used to improve the lightning protection of transmission lines in the mountain areas. It takes the differences of various factors such as distribution of lightning activity, topography, tower structure, and lightning accident into consideration with the basic steps shown as below:

1. Parametric statistics
2. Analysis of the lightning protection performance
3. Risk assessment
4. Lightning protection configuration of the transmission line and program formulation
5. Technical and economic evaluation of the rehabilitation programs.

52.3 Parametric Statistic

This chapter selects the 110 kV Transmission Line A subject to trip-out mostly in the region as an example to transform lightning protection. Line A of 28.178 km in length, whose transmission lines are mainly located at the mountains, was put into use on January 23, 2003. Sixteen cases of lightning accident occurred from 2003 to 2005 on this line, so its lightning trip-out rate was 28.39 times/(100 km·year), much higher than the specified value 0.97. Relevant records show that effective measures are not taken to reduce the lightning trip rate except that the lightning rods of tower top were installed on the breakdown towers.

It is important to master the lightning distribution of the transmission line corridor and obtain the actual parameters of lightning so as to achieve the reliability of differentiated lightning protection. The ground flash density and the amplitude distribution of lightning current are the foremost parameters [8].

Table 52.1 Ground flash density

	LLS					
	06	07	08	09	10	11
GFD (f/km^2·y)	5.01	3.87	5.72	3.99	6.15	6.04
	Average:5.13506					

Table 52.2 Parameter distribution of the lightning current amplitude curve

Year	IEEE criterion (P=1/[1+(I/x)^y]	
	x	*y*
2006	32.7567	2.9747
2007	28.4083	3.3864
2008	28.4958	2.7196
2009	28.8115	2.7606
2010	27.7463	2.6541
2011	24.6202	2.6587

Ground flash density (GFD) is available according to the lightning location system (LLS) as shown in Table 52.1:

The amplitude distribution curves of lightning current during 2006 to 2011 can be obtained according to lightning parametric statistics. These curves are substituted to the following Formula (52.1).

In Formula (52.1), *I* is the amplitude of the lightning current; *x* is the mean current; in other words, the probability of lightning current exceeding *x* is 50%; *y* reflects the change in the probability curve; *p* (Ip>I) represents that the probability of lightning current is larger than I. The amplitude of lightning current from 2006 to 2011 is fitted to curves. The distribution parameters *x* and *y* derived from the fitted curves are shown in Table 52.2:

$$p(\mathrm{I}_p > \mathrm{I}) = \frac{1}{1+\left(\mathrm{I}/\mathrm{x}\right)^{\mathrm{y}}}. \tag{52.1}$$

The average probability distribution of each corresponding amplitude point can be obtained, as shown in Table 52.3, according to the calculation of the average probability distribution of lightning current amplitude during the 6 years.

The data of Table 52.3 can be fitted into the amplitude distribution curve of the average lightning current. Formula (52.2) is calculated as below:

$$p(\mathrm{I}_p > \mathrm{I}) = \frac{1}{1+\left(\mathrm{I}/28.9886\right)^{2.8619}}. \tag{52.2}$$

Table 52.3 Parameter distribution of average probability of lightning current from 2006 to 2011

I	10	20	30	40	50	60	70	80	90	100
p (Ip>I)	0.984	0.737	0.472	0.283	0.174	0.113	0.080	0.056	0.040	0.031
I	110	120	130	140	150	160	170	180	190	110
p (Ip>I)	0.025	0.019	0.016	0.014	0.010	0.007	0.008	0.004	0.005	0.003

52.4 Analysis of Lightning Protection Performance and Risk Assessment

The lightning withstanding level (kA) and the lightning trip-out rate (times/100 km·year) are two important indexes of characterizing the lightning protection performance of transmission lines. The trip-out rate by back-flashover is calculated based on the electromagnetic transient program EMTP while the shielding failures are based on the approach of the pilot development model [9].The lightning trip-out rate of lines can be roughly estimated based on the parameters about line and terrain. While compared with the actual data, the result can be regarded as the evaluation criteria of lightning protection performance.

First, a standard for risk assessment has to be established. According to provisions of Article 89 of *110 (66) kV500 kV Overhead Transmission Line Management Practices* as released by the National Grid Company Sets [10], the lightning trip rate of each voltage level normalized to 40 thunderstorm days should not exceed the following indicators: 110 (66) kV; 0.525 (times/100 km·a). If converted to the annual average ground flash density in the region (5.13506 f/km^2·a), the corresponding index should be: 110 (66) kV; 0.970 (times/100 km·a).

Upon calculation of the back-flashover and shielding trip rate of each tower, the risk level of lightning flashover can be classified. The criterion of classification used for the risk assessment of lightning flashover of overhead transmission lines in the region is shown in Table 52.4.

In Table 52.4, *S* represents the average lightning trip-out rate. *S* of Line *A* is 2.452 (times/100 km·a) which is larger than the prescriptive value. The lightning protection performance will be divided into four grades (A, B, C, and D) according to the standard in Table 52.4. Upon calculation of the lightning trip-out rate of each tower across Line *A*, the risk grade of each tower is shown in Fig. 52.1.

The grade distribution of towers is based on the calculation results as shown in Table 52.5.

Table 52.4 Standard for assessment of lightning risk

Risk assessment level	A	B	C	D
Trip rate	$P<0.5\cdot S$	$0.5\cdot S\leq P<1.0\cdot S$	$1.0\cdot S\leq P<1.5\cdot S$	$P\geq 1.5\cdot S$

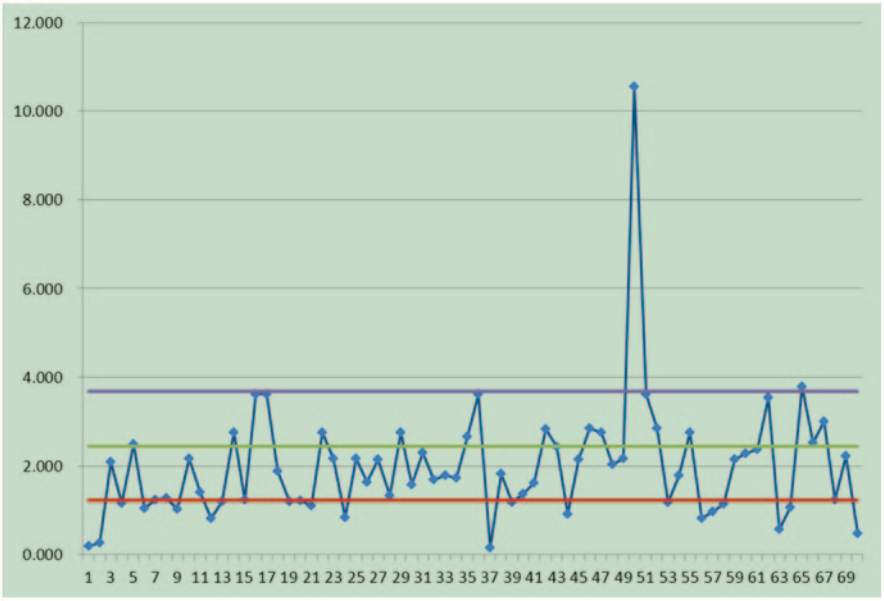


Fig. 52.1 Results of the trip rate and lightning risk assessment

Table 52.5 Distribution of lightning risk grade

Risk assessment level	A	B	C	D
Range of trip rate (times/100 km·y)	0~1.226	1.226~2.452	2.452~3.678	≥3.678
Amount of towers	21	30	17	2
Percentage (%)	30%	42.86%	24.29%	2.85%

52.5 Program of Lightning Protection Transformation

The flow chart of lightning protection transformation is shown as below (Fig. 52.2):

As the transmission line is already built, the reduction of resistance and installing line arresters are selected as the main measures and installing controllable lightning rod as an auxiliary measure. As to the measurement of resistance reduction, it suggests that the relevant unit should retest the earth resistance and determine the target of grounding transformation in combination with the results of the risk assessment [11]. As to the installing line arrester, the choices of the installation phase and the number of line arresters refer to the terrain of the tower and the probability of shielding flashover [12].

The scope and sequence of the towers for transformation are determined as below on the basis of the lightning trip records of this line and the theoretical calculation of the lightning trip rate:

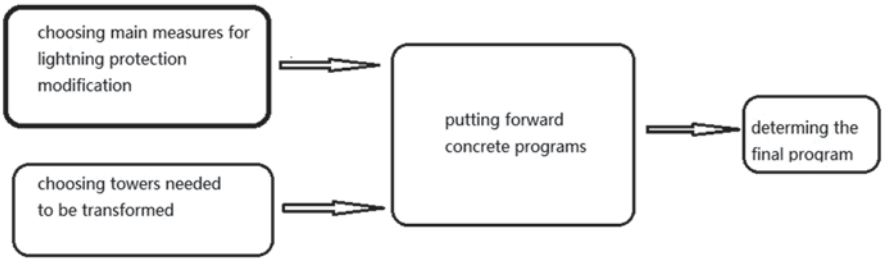


Fig. 52.2 The flow chart of lightning protection transformation

1. Towers that have been subject to lightning
2. Towers that are evaluated as grade D
3. Towers that are evaluated as grade C.

There are 16 faults available in two years on Line *A*. The 17th tower has been tripped 4 times; NO. 21, 30, 61 towers all have been tripped 2 times; the other six towers have been tripped 1 time. The following protective measures should be taken based on the reasons for the tower fault:

1. The measure of reducing grounding resistance will be taken for No. 17, 61, 18, 35, and 60 towers
2. Line arresters will be installed at No. 21, 30, and 61 towers (the 17th tower has been installed)
3. The Controllable lightning rods or line lightning arresters will be installed on both sides of the wire for No. 18, 35, 60, and 63 towers (No. 22 and 31 towers have been installed).The impact of economic factors should be taken into consideration in a specific program.

The result of risk assessment shows that there are 17 towers of grade C and two of grade D. All of these towers of grades C and D are weak. There are three towers (NO.17, 22, and 35) of grade C that have been subject to lightning with three (NO. 29, 46, and 47) installed with the line arresters. The 50th tower of grade D has been installed with the line arrester. In other words, there are 12 weak towers in total.

There are 7 towers without the arrester subject to lightning, 11 towers of grade C without arrester not subject to lightning and only one of grade D; in this sense, 19 towers have to be transformed. The specific programs are shown in Table 52.6.

In Program 1, the towers which are struck by lightning more than one time will be instilled with the line arrester; in Program 2, the towers hic hare stuck by lightning will be instilled with the line arrester; in Program 3, the towers of grade D which are stuck by lightning will be instilled with the line arrester; in Program 4, the towers of grades C and D, which are stuck by lightning will be instilled with the line arrester. Other towers which need to be transformed will be instilled with the controllable lightning rods according to their actual situations. Each tower will be installed with two lightning arresters or rods. The price of a line arrester is about RMB7000 and a lightning rod of about RMB7500.

Table 52.6 Schemes of modification of the lightning protection

Program	Lightning rod		Line arrester		Expected trip rate (f/100km·a)	Percentage of better than the standard	Expense (thousand yuan)
1	16	32	3	6	1.043	–	28.2
2	12	24	7	14	1.009	–	27.8
3	11	22	8	16	0.997	–	27.7
4	0	0	19	38	0.816	16.0%	26.6

It is shown in Table 52.6 that only Program 4 complies with the standard; therefore, the final concrete measures are shown as below:

1. The measure of resistance reduction will be taken for NO. 17, 61, 18, 35, and 60 towers
2. Two line arresters will be instilled on NO. 5, 14, 16, 18, 21, 29, 30, 35, 36, 42, 51, 52, 55, 60, 62, 63, 65, 66, and 67 towers.

Conclusion: The existing lightning fault trip records and the risk assessment result of each tower of Line A are regarded as the criteria for the lightning protection transformation. According to the results, it selects the towers that need to be transformed. Then the lightning protection performance before and after reconstruction will be compared. The results show that the final program has improved the lightning protection performance of Line A, and it has greatly reduced the risk of lightning flashover of the line.

52.6 Conclusion

The main content of this chapter concerns the application of differentiated lightning protection technology to power transmission line in the mountain areas. It uses differentiated lightning protection technology to transform the lightning protection of Line A which features the highest risk of lightning flashover as to 110 kV transmission lines in the region. It implements the steps of this technology and puts forward a reform program finally.These results show that the scheme has greatly improved the lightning protection performance of Line A. Through this study, the following conclusions can be obtained:

1. In this chapter, a 110 kV transmission line, which has a serious lighting fault, has been selected to carry out the risk assessment of lightning flashover and calculate the lightning trip rate of each tower. What is more, compared to the risk level classification criteria and results of the trip rate, there comes the risk assessment level of each tower. As the result of risk assessment, the towers of grades C and D account for 27.14%
2. According to the actual situation, the main reform measures using the dynasty blocking transformation and installation of line arresters have been taken;

the installation of controlling discharge lightning rod has been regarded as an assisted measure. According to the results of the risk assessment and the fault statistics, it will determine the towers which are going to be changed. Reasonable choice of specific measures will be taken based on the actual situation, and it eventually formulates a self rehabilitation program to meet the requirements by comparing the effect and economics

3. As to the lightning protection transformation of the transmission line, the principle of transformation, the calculation method, the reference criterion, and the ultimate goal have nothing to do with the voltage level; therefore, the method can also apply to transmission lines of different voltage levels in mountain areas, such as 220 kV and 330 kV. At the same time, it provides a very effective method for the similar engineering problems in fact.

References

1. State Grid Electric Power Research Institute. Research on power transmission and transformation equipment lightning protection technology and strategy [R]. Wuhan: State Grid Electric Power Research Institute; 2009 (In Chinese).
2. Wang Z-H. Calculation of shielding failure trip-out rate for transmission lines in mountain areas [J]. High Voltage Engineering. 1992;4(02):28–33 (In Chinese).
3. Xu Z-Y. Study on lightning protection performance of high voltage transmission line in mountainous area and countermeasures [D]. Nanning: Guangxi University; 2008 (In Chinese).
4. Zhou Y-X, Lu B, Yan F-L, et al. Shielding failure flash-over rate under conditions of complex landscape in a mountainous area [J]. High Voltage Engineering. 2007;33(06):45–9 (In Chinese).
5. Gu S-Q, Chen J-H, Chen W-J, et al. Evaluation method of the time-space difference of lightning protection performance of transmission lines [J]. High Voltage Engineering. 2009;35(02):294–8 (In Chinese).
6. Chen J-H, Lu J, Qian Z-Y, et al. Differentiation technology and strategy of lightning protection for transmission lines [J]. High Voltage Engineering. 2009;35(12):2891–902 (In Chinese).
7. Zhao C, Ruan J-J, Li X-L, et al. Technology and economy evaluation of comprehensive transmission line lightning protection measures [J]. High Voltage Engineering. 2011;37(02):290–7 (In Chinese).
8. Chen J-H, Feng W-X, Wang H-T, et al. Statistical method of lightning paraments [J]. High Voltage Engineering. 2007;33(10):6–10 (In Chinese).
9. Zeng R, Geng Y-N, Li Yu, et al. Lightning shielding failure model of transmission line based on leader progress model [J]. High Voltage Engineering. 2008;34(10):2041–6 (In Chinese).
10. State Grid Corporation of China. Operation specifications for 110(66)~500 kV overhead transmission lines [S]. Beijing: State Grid Corporation of China; 2005 (In Chinese).
11. Li C. Research on reducing tower ground resistance of Yamaiwa region by using tower foundation [D]. Changsha: Changsha University of Science and Technology; 2008 (In Chinese).
12. Chen X-G, Yuan T, et al. Application of line ZnO arrester on lightning protection of transmission line [J]. High Voltage Engineering. 2003;29(12):17–9 (In Chinese).

Chapter 53
The Selection Strategy of Form-Finding Methods for Transmission Lines Based on ANSYS

Ying Jin, Changsheng Liu, Yunyun Xie and Linghao Zhang

Abstract The aim of this chapter is to obtain the selection strategy of form-finding methods for the initial state of transmission lines based on ANSYS. First, this chapter introduces the theory and the main characteristic of four kinds of form-finding methods for transmission lines, which are the "V"-shape method, the little elastic modulus method, the direct iteration method, and the across-the-lowest-point method. Second, in order to get the selection strategy of these four methods, design a group of calculation models. Then, this chapter organizes simulation results on the basis of different methods, and makes further analyses of these data. The selection strategy proposed in this chapter can provide a convenient way to choose the proper form-finding method for the initial state of transmission lines based on ANSYS.

Keywords Selection strategy · ANSYS · Form-finding · Transmission lines

53.1 Introduction

ANSYS is not only powerful, but also easy to learn. With the help of ANSYS software, studying the electric power system becomes much easier [1, 2]. For the simulation problem of transmission lines in ANSYS, the form-finding method for the initial state of transmission lines plays an important role because of its basic and irreplaceable effect [3]. Although there exist several form-finding methods for transmission lines, no one put forward any selection strategy of existing form-finding methods based on ANSYS. Therefore, this chapter aims at offering a useful selection strategy of different form-finding methods which can provide the most suitable form-finding methods for each calculation condition.

Four common form-finding methods for transmission lines based on ANSYS introduced in this chapter are the "V"-shape method, the little elastic modulus method, the direct iteration method, and the across-the-lowest-point method. WH Zhang et al. proposed the direct iteration method (called as Method I below) which

Y. Xie (✉) · Y. Jin · C. Liu · L. Zhang
School of Automation, Nanjing University of Science and Technology, 210094 Nanjing, China
e-mail: xnjust@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_53

proceeds nonlinear iteration under the effect of deadweight [4–7]. YB Lu et al. proposed the "V"-shape method (called as Method II below), and its initial geometric model is designed like the letter "V" [1, 2, 8–10]. YZ Jia et al. proposed the little elastic modulus method (called as Method III below) which sets the smaller elastic modulus value as the initial value [11–13]. Ying Jin proposed the across-the-lowest-point method (called as Method IV below) which stops the iteration when the deformed model is across the lowest point of transmission lines [14–15].

Because the accuracy of each method differs in distinct calculation conditions [14], the existing problem of form-finding methods is that which form-finding method is the best choice according to the specific condition. Hence, it is necessary to create a selection strategy of form-finding methods for transmission lines.

53.2 Four Conductor Form-finding Methods

53.2.1 Method I

The theory of Method I is that the displacement of geometric model is simply generated by gravity. Usually, the horizontal tension value is set as convergence value. Although the operation of this method is not complicated; Method I needs multiple iterations which might spend much time [4, 7, 15, 16].

53.2.2 Method II

The main characteristic of Method II is to build a geometric model in the shape of letter "V." Furthermore, the length of "V"-shape model should be equal to the theoretical value of catenary which is calculated by catenary equation [12, 13]before. The significant advantage of this method is that the initial form of transmission lines can be found after iteration [1, 2].

53.2.3 Method III

The theory of Method III is that the displacement of geometric model is not simply produced by gravity which is opposite to that of the direct iteration method. Therefore, in this method, by means of gravity and adjusting the elastic modulus value, the initial state of transmission lines still can be found [8, 11, 17]. However, adjusting the elastic modulus depends on personal experience which might produce big error.

53.2.4 Method IV

The theory of Method IV is that consider the deformed model which is across the lowest point calculated by catenary equation [12, 13] as the initial form of transmission lines [15]. Similarly, this method needs adjusting the elastic modulus value, in order to make the geometric model pass through the lowest point.

53.3 Examples

53.3.1 Design of Examples

The type of transmission conductor is LGJ-300/40 [12]. There are three kinds of calculation models, and each model has three calculation conditions which are described in Table 53.1.

53.3.2 Results

This chapter discusses three types of data: the length of catenary, the sag of semispan, and the stress of left suspension point. Besides, the theoretical value of each result is calculated by catenary equation [12, 13]. Moreover, if there is height difference within a span, errors of Method II are unacceptable [1, 2, 14]. So, in Model II and Model III, we do not consider about Method II. Simulation results and theoretical values of each calculation condition are shown in Tables 53.2 and 53.3, respectively.

According to the errors between theoretical values and simulation results, the scatter diagrams of each model are shown in Figs. 53.1, 53.2, 53.3:

Table 53.1 Calculation conditions

Calculation conditions		Descriptions of calculation conditions
Model I	S-1–1	No height difference; span is 50m
	S-1–2	No height difference; span is 450m
Model II	S-2–1	Height difference is 2 m; span is 200 m
	S-2–2	Height difference is 10 m; span is 200 m
Model III	S-3–1	Height difference is 30 m; span is 200 m
	S-3–2	Height difference is 50 m; span is 200 m

Table 53.2 Simulation results

Calculation conditions		Model I		Model II		Model III	
		S-1–1	S-1–3	S-2–1	S-2–3	S-3–1	S-3–3
Catenary length (m)	Method I	150.05	451.15	200.17	200.41	202.40	206.31
	Method II	150.06	451.41				
	Method III	150.07	451.11	200.20	200.44	202.43	206.35
	Method IV	150.06	451.83	200.20	200.40	202.42	206.33
Sag of semispan (m)	Method I	1.73	13.94	3.44	3.45	3.50	3.59
	Method II	1.75	15.45				
	Method III	1.96	13.73	3.80	3.81	3.86	3.97
	Method IV	1.78	17.58	3.75	3.74	3.76	3.76
Stress of the left suspension point (N/mm^2)	Method I	52.54	58.32	47.64	47.56	47.51	47.76
	Method II	52.24	53.52				
	Method III	47.08	60.93	43.24	43.13	43.04	43.17
	Method IV	51.81	47.70	43.79	43.79	44.17	45.29

Table 53.3 Theoretical values

Calculation conditions		Catenary length (m)	Sag of semispan (m)	Stress of the left suspension point (N/mm^2)
Model I	S-1–1	150.23	1.75	51.91
	S-1–2	457.33	17.58	47.87
Model II	S-2–1	200.75	3.75	43.85
	S-2–2	200.75	3.75	43.81
Model III	S-3–1	200.74	3.75	44.30
	S-3–2	200.71	3.76	45.63

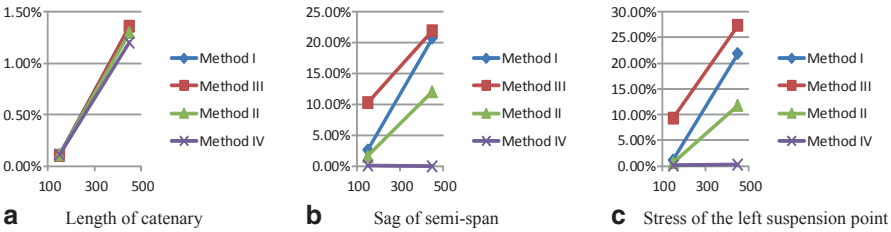


Fig. 53.1 Scatter diagram of Model I

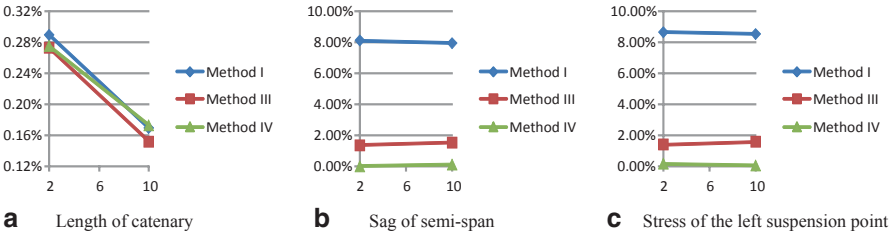


Fig. 53.2 Scatter diagram of Model II

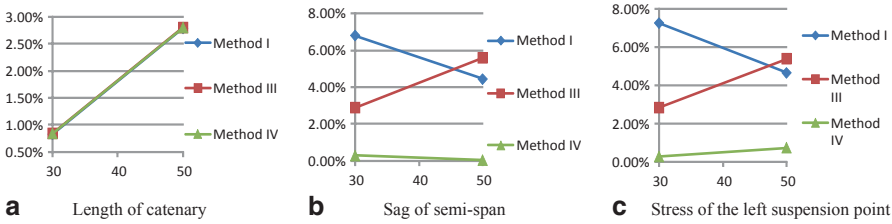


Fig. 53.3 Scatter diagram of Model III

53.3.3 Discussion of Results

The results indicate that the accuracy of these four form-finding methods is the same when they are used in the identical calculation conditions. Because all the error curves coincide with each other in the scatter diagram of catenary length, the accuracy is always the same; no matter we use which form-finding method in specific calculation conditions.

On aspects of sag of semispan and stress of the suspension point, we believe that the accuracy of Method IV is the highest while the accuracy of other form-finding methods relies on calculation conditions. Specifically, if the span distance is short without height difference, the errors of Method I and Method II are acceptable; if the span distance is long, their errors are unacceptable. Moreover, in Fig. 2, curves of Method I and Method III are both paralleled to horizontal axis which means that the height difference does not affect the accuracy of these two methods; but the accuracy of Method III is much higher than that of Method I. Besides, in Fig. 3, with the increment of height difference, the accuracy of Method I is increasing; however, the accuracy of Method III is decreasing.

The reasons for the great change of accuracy of Method I and Method III are as follows. First, they do not have unified and standard criterion of iteration stopping, and usually the iteration cannot stop until errors of simulation results meet the requirement set by researchers. Second, the initial elastic modulus value in Method III is set according to personal experience, which might exert harmful influence on simulation results.

53.4 Generation of Strategy

53.4.1 Analysis of Accuracy

Use numbers 5–1 to express the accuracy level of these form-finding methods, and the accuracy level of each method is shown in Table 53.4.

53.4.2 Analysis of Difficulty

According to the flow diagrams given above, operational difficulty levels of each form-finding method can be described as numbers: 5, 4, 3, 2, and 1, from the easiest level to the most difficult level. And the difficulty analysis of each method is given in Table 53.5.

53.4.3 Selection Strategy

By combining the analyses of accuracy and difficulty, this chapter proposes a selection strategy of conductor form-finding method for different data types and calculation conditions. The selection strategy is given in Tables 53.6.

Table 53.4 Accuracy analysis of three models

Calculation conditions		Direct iteration method	"V"-shape method	Little elasticity modulus method	Across-the-lowest-point method
S-1–1 S-2–1 S-3–1	Catenary length	5+5+5	5	5+5+5	5+5+5
	Sag of semispan	4+3+3	4	1+4+4	5+5+5
	Stress of point	4+3+2	5	4+4+4	5+5+5
S-1–2 S-2–2 S-3–2	Catenary length	4+5+4	4	4+5+4	4+5+4
	Sag of semispan	1+3+3	1	1+4+3	5+5+5
	Stress of point	1+3+3	1	1+4+3	5+5+5

Table 53.5 Difficulty analysis of operations

Direct iteration method	"V"-shape method	Little elasticity modulus method	Across-the-lowest-point method
3	5	2	1

Table 53.6 Selection strategy

Data type			Transmission conductor form-finding methods
Model I	Catenary length	Larger span	Across-the-lowest-point method
		Smaller span	"V"-shape method
	Sag of semispan	Larger span	Across-the-lowest-point method
		Smaller span	"V"-shape method
	Stress of suspension point	Larger span	Across-the-lowest-point method
		Smaller span	"V"-shape method
Model II	Catenary length sag of semispan stress of suspension point		Direct iteration method little elasticity modulus method Little elasticity modulus method
Model III	Catenary length		Direct iteration method Across-the-lowest-point method Across-the-lowest-point method
	Sag of semispan		
	Stress of suspension point		

53.5 Conclusion

This chapter does research on four kinds of form-finding methods for transmission lines: "V"-shape method, little elastic modulus method, direct iteration method, and across-the-lowest-point method. On the foundation of theoretical analysis, three kinds of calculation models are designed. After simulation of ANSYS, this chapter proposes a selection strategy of conductor form-finding methods which is of practical value.

References

1. B Li, PH Gong. Self-study manual ANSYS14 of finite element analysis. Beijing: Posts and Telecom Press; 2013. pp. 239–50 (In Chinese).
2. YB Lu. Analysis of ice-shedding and wire-breaking in transmission line system. Hangzhou: Zhejiang University; 2010 (In Chinese).
3. W Kong, MW Zhu, H Fu. Research on form-finding method for transmission lines in ANSYS. J Northeast Electr Power Univ. 2012;31(5/6):64–7. (In Chinese).
4. SL Chen, MG Cao. Analysis of form-finding of overhead transmission lines and grounding line based on the application of ANSYS. Jiangxi Electric Power. 2007;31(6):7–15. (In Chinese).
5. TX Shao. Mechanical calculation of overhead transmission lines. 2nd edn. Beijing: China Electric Power Press;2003. pp. 181–90 (In Chinese).
6. Y Wang. Analysis and realization of form-finding of cable-membrane structure by finite element method. J Liaoning Tech Univ. 2011;24(4):104–7. (In Chinese).
7. WH Zhang, JB Yu. Form-finding analysis of overhead transmission line based on ANSYS. Electric Power Constr. 2012;33(2):32–5. (In Chinese).
8. YZ Jia, RP Liu, ZQ Li. Original state for overhead icing transmission line. Water Resour Power. 2011;29(1):148–50. (In Chinese).
9. D Veenendaal, P Block. Overview and comparison of structural form-finding methods for general networks. Solids Struct. 2012;49:3741–53.
10. DS Zhang. Design manual of high voltage transmission lines in electric power engineering. 2nd edn. Beijing: China Electric Power Press;2002. pp. 177–98. (In Chinese).
11. YB Lu, WJ Lou, HL Li. Process analysis of non-uniform ice shedding of transmission lines. Vib Shock. 2010;29(9):47–50. (In Chinese).
12. SM Meng, LP Shan. Form-finding in dynamics analysis of transmission line. Adv Power Syst Hydroelectr Eng. 2009;25(10):43–7. (In Chinese).
13. Q Yang, CM Li. Form-finding and simulation of cable structure in ANSYS. Inf Technol Civ Eng Archit. 2010;2(4):61–5. (In Chinese).
14. Y Jin. Research on comparison of form-finding methods for transmission lines based on ANSYS. Proceedings of 2014 international conference of electrical and electronics engineering. Unpublished (2014)
15. Y Jin. Research on form-findings of conductor with consideration for the lowest point in transmission lines. Proceedings of the 2nd international conference on structural health monitoring and integrity management. Unpublished (2014)
16. WH Zhang. Numerical simulation and experimental research on vibration of iced overhead conductors. Beijing: North China Electric Power University; 2012. (In Chinese).
17. ZC Xia, L Li, ZP Liang et al. Dynamic response of transmission tower with ruptured wires. Vib Shock. 2007;26(11):45–9. (In Chinese).

Part V
Theory and New Technology

Chapter 54
An Artificial Bee Colony Algorithm Guided by Lévy Flights Disturbance Strategy for Global Optimization

Wei Ma, Zhengxing Sun, Junlou Li, Mofei Song, Xufeng Lang and Cheng Le

Abstract In order to solve the problems of the slow convergence speed, low precision and easy trapping in local optimal solutions in an artificial bee colony (ABC) algorithm, a novel modified ABC algorithm based on the Lévy flights disturbance mechanism is proposed in this chapter. It attempts to increase the exploration efficiency of the solution space for global optimization. The modifications focus on the solution construction phase of the artificial bee colony algorithm. In addition, to further balance the search processes of exploration and exploitation, the modification forms of the onlookers and scouts search strategy is proposed in this chapter. It could avoid local optimum. And it also could greatly improve convergence speed and solution precision on the basis of keeping strong global optimization performance of the proposed algorithm. Simulation experiment results based on typical benchmark functions show that the proposed algorithm is more effective to avoid premature convergence and to improve solution precision than some other ABCs and several state-of-the-art algorithms.

Keywords Cuckoo search (CS) algorithm · Artificial bee colony algorithm · Lévy flights · Global optimization · Disturbance strategy

W. Ma (✉) · J. Li
Nanjing Institute of Tourism and Hospitality, 211100 Nanjing, China
e-mail: mawei@smail.nju.edu.cn

W. Ma · Z. Sun · · M. Song · · X. Lang
State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, 210093 Nanjing, China
e-mail: mawei@smail.nju.edu.cn

C. Le
Department of Computer Science and Engineering, Huaian College of Information Technology, 223003 Huaian, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_54

54.1 Introduction

Artificial bee colony (ABC) [1] algorithm was introduced by Karaboga in 2005. It is a stochastic optimization algorithm which simulates the intelligent foraging behavior of honey bees. Subsequently, Karaboga and Basturk compared the performance of the ABC algorithm with that of other well-known modern heuristic algorithms such as GA, PSO, differential evolution (DE), and particle swarm optimization evolutionary algorithm (PS-EA) on unconstrained and constrained problems [2]. The study showed that the performance of the ABC algorithm was better than or similar to those of other population-based algorithms with the advantage of employing fewer control parameters. The comparison results showed that ABC algorithm was effective. The algorithm has not only been widely applied to the areas of discrete optimization, but also been applied for solving a variety of real-life benchmark problems [3], such as function optimization problems.

Two research strategies on further exploring the performance of continuous space function optimization in the ABC algorithm have been adopted. The first research strategy is that the search mechanism based on the basic ABC algorithm is modified partly [4–11]. Zhu and Kwong proposed best-guided artificial bee colony (GABC) algorithm [4] which utilizes global best solution to improve the exploitation of the solution search equation. The second research strategy is to combine ABC with other optimization algorithms [12–14], such as improved artificial bee colony algorithm (IABC) [12], Rosenbrock artificial bee colony algorithm (RABC) [13], hybrid harmony search with artificial bee colony algorithm (HHSABC) [14], etc. Although the search performance could be improved by the combination of ABC and other optimization algorithms, the original intention of simulating biological characteristics in the basic ABC algorithm has been changed. And the calculation time complexity of the algorithm has been increased accordingly. However, the proposed algorithm combines the characteristics of Lévy flights disturbance mechanism with the modification strategy of search mechanisms of onlookers and scouts based on the ABC algorithm. It is a new evolutionary optimization algorithm which based on these two research strategies.

The strategy of Lévy flights disturbance mechanism was inspired by the cuckoo search (CS) algorithm, which was a new intelligent optimization algorithm proposed by Yang from Cambridge University and Deb from Raman Engineering University in 2009 [15]. Although the ABC algorithm has better global search ability, its relative shortage of local search capability and solution precision still needs to be enhanced. According to the study of real bee colony activities, we found that the flight path of bee colony in the process of foraging possesses the characteristics of Lévy flights. Inspired by the CS algorithm, the Lévy fights artificial bee colony algorithm (LFABC) based on the Lévy flights disturbance strategy has been proposed in which the search mechanism of onlookers and scouts is modified. It can overcome the shortcomings of the ABC algorithm. Our numerical experimental results show that the LFABC algorithm is very satisfactory with fast convergence speed, success rate, high solution precision, and searching efficiency.

The rest of this chapter is organized as follows. The proposed LFABC algorithm, which is based on the Lévy flights disturbance strategy, is presented in Sect. 2. The experimental results conducted by benchmark functions are presented and discussed in Sect. 3. Finally, the conclusion is drawn in Sect. 4.

54.2 The ABC Based on the Strategy of Lévy Flights Disturbance

54.2.1 Lévy Flights Search Mechanism

The CS algorithm forms theory by the way cuckoo searches nest. This search algorithm is designed based on Lévy flights search mechanism. The simple ideal rules of the algorithm are defined as follows: Every cuckoo lays one egg at a time, and selects randomly parasitized nests for hatching. The best one would be reserved to the next generation among the selected parasitized nests. The number of available parasitized nests is fixed, and the eggs laid by a cuckoo are discovered by the host birds with certain probability. According to the rules defined above, the host birds could either throw the eggs away or abandon the nests and build a new one.

The path and position cuckoo searched for could be expressed as Eq. (54.1). In order to generate new solutions x_i^{t+1}, a Lévy flight is performed as follows:

$$x_i^{t+1} = x_i^t + \alpha \oplus \mathrm{Levy}(\lambda), \tag{54.1}$$

where, x_i^t denotes ith nest locating in the position of tth generation bird nest. $\alpha > 0$ is the step size. In most cases, take $\alpha = 1$. The Eq. (54.1) is an essentially random walk equation. Normally, a random walk is a Markov chain, and its next status depends on the current location of the position (the first term in the above equation) and transition probabilities (the second item). The product $\oplus$ means entrywise multiplication, while $\mathrm{Levy}(\lambda)$ is a random search path. The Lévy flights essentially provide a random walk, while the random steps length are drawn from a Lévy distribution for large steps, which could be seen from Eq. (54.2)

$$\mathrm{L\emptyset vy} \sim u = t^{-\lambda},\ 1 < \lambda \le 3. \tag{54.2}$$

54.2.2 Lévy Flights Disturbance Strategy

The Lévy flights search mechanism is adopted to enlarge searching a range of solution space, enhance the search diversity of bee colony in the intelligent optimization algorithm. In this way, it makes the algorithm escape from the local minima easily.

The random steps are generated by adopting Lévy flights in the CS algorithm. The step size is always changeable. In the process of searching, the larger the step

size is, the easier the global optimal solution gets. The solution precision is reduced accordingly, and it is easy to produce unstable vibration situation. If the step size is smaller, the search speed is slowing down, and the solution precision is enhancing significantly. Therefore, the step produced by Lévy flights is random and ergodic. The drawback is that, adaptive capacity is not enough. While the ABC algorithm could cover the shortage, neighborhood search mechanism of the onlookers is adopted to balance the search processes of exploration and exploitation. Therefore, the global search ability and the local solution accuracy of the ABC algorithm are enhanced.

The construction of Lévy flights step disturbance strategy is inspired by the CS algorithm. The definition of disturbance step is calculated by Eq. (54.2) as follows:

$$s_i = \frac{u}{|v|^{1/\beta}} \tag{54.3}$$

$$u \sim N(0, \delta_u^2), v \sim N(0, \delta_v^2) \tag{54.4}$$

$$\delta_u = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \delta_v = 1 \tag{54.5}$$

$$\text{step}_i = \tau \cdot s_i \cdot x_{\text{best}}^j, \tag{54.6}$$

where $i \in \{1, 2, \ldots, SN\}$, $j \in \{1, 2, \ldots, D\}$, u, v is normal distribution, $\beta = 1.5, \tau = 0.01$, the position of employed bees is updated by using the original update policy based on step disturbance which is defined as Eq. (54.7):

$$v_{i,j} = v_{i,j} \bullet \text{step}_i. \tag{54.7}$$

54.2.3 Modified ABC Search Strategy

54.2.3.1 Search Mechanism Based on the Onlookers

The search strategy of the basic ABC algorithm has been modified in this chapter. The employed bees share the nectar information with the onlookers by waggle dance, when the search based on Lévy flights disturbance strategy has been finished. The onlookers select food sources by the nectar information provided by the employed bees in certain probability in the basic ABC algorithm. The higher the nectar amount is obtained, the greater probability it has to attract onlookers. The attracted onlookers search in the neighborhood of food sources. In order to utilize the search results got by the Lévy flights disturbance mechanism, this chapter makes the food sources below the probability by the way of Eq. (54.8) to attract

the onlookers searching toward global optimal food sources. In this way, the search diversity of bee colony is enhanced.

$$v_{i,j} = x_{i,j} + \varphi_{i,j}(x_{i,j} - x_{best,j}), \tag{54.8}$$

where $\varphi_{i,j} = \text{rand}[-1,1], x_{\text{best},j}$ denotes the value of j-dimensional vector corresponding to the position of the global optimal nectar.

54.2.3.2 Search Mechanism Based on the Scouts

For the purpose of making ABC algorithm escape from the local optima easily, this chapter is inspired by the modified strategy used in best-so-far ABC and GABC algorithms, to further modify the search mechanism by using Eq. (54.9) to replace the random search mechanism used in the basic ABC algorithm. And linear inertia weight is imported to guide the global search in this chapter. The new strategy for position renewal of employed bees is expressed by Eq. (54.9):

$$v_{i,j} = x_{i,j} + \varphi_{i,j} \bullet w \bullet x_{i,j}, \tag{54.9}$$

where $\varphi_{i,j} = \text{rand}[-1,1], w$ is linear inertia weight. It is calculated by Eq. (54.10) as follows, where $w_{\max} = 1, w_{\min} = 0.2$:

$$w = w_{\min} + \frac{\text{iteration}}{MCN} \bullet (w_{\max} - w_{\min}). \tag{54.10}$$

54.2.4 Objective Value-Based Selection Optimization

In the basic ABC algorithm, the selection of quality food sources is based upon the fitness value which is used to evaluate the nectar amounts. And the fitness value could be obtained from Eq. (54.11). If the fitness value of new food sources is higher than the previous food source, then it replaces the previous one. However, from Eq. (54.11), the corresponding fitness value approaches 1 infinitely, when the value of the objective function is infinitely close to 0. And when the value of the objective function is smaller than certain magnitude, it is very difficult to distinguish the size of the fitness value. To solve this problem, the chapter is inspired [5] to replace the fitness value with the objective value-based directly. In order to facilitate the comparison experiment, two algorithms are presented. One of the algorithms is based on the evaluation of the fitness value, named LFABC1. The other one is based on the value of objective function, named LFABC:

$$fit_i = \begin{cases} \dfrac{1}{1+f_i}, & f_i \geq 0 \\ 1+|f_i|, & f_i < 0 \end{cases}. \tag{54.11}$$

54.2.5 The Steps of LFABC Algorithm

The detailed steps of the LFABC algorithm are given below:

step1: Parameter settings: Set the population of the bee colony *m*, set the counter of bee colony steps $t = 0$, the maximum number of bee colony steps is *MCN* and set the maximum limited times of staying food source *limit*

step2: Initialization: According to Eq. $x_{i,j} = x_j^{lb} + \text{rand}(0,1)(x_j^{ub} - x_j^{lb})$ in the search area, initialize the solutions $X_i (i = 1, 2, \cdots, m)$, evaluate the values of objective function, the fitness value of the top *m/2* as food sources, and correspond to *m/2* employed bees. Initialize vector $Bas(i) = 0$ which records the cycle number of each food source the employed bees remain

step3: The employed bees searching phase: each employed bee searches the food sources by using Eq. $v_{i,j} = x_{i,j} + \varphi_{i,j}(x_{i,j} - x_{k,j})$, move a certain step according to Lévy flights disturbance by means of Eq. (54.7), and gets new position of food source V_i. The value of objective function $y = f_{opt}^{-}(V_i)$ is calculated, and a better food source is selected as the next search position between the new food source $V_i (i = 1, 2, \cdots, m_e)$ based on the evaluation and the previous one $X_i (i = 1, 2, \cdots, m_e)$; if $f_{opt}^{-}(V_i) < f_{opt}^{-}(X_i)$, set $X_i = V_i, Bas(i) = 0$, otherwise, set $Bas(i) = Bas(i) + 1$

step4: Recruitment selection mechanism: After the employed bees finish searching in the neighborhood of all food sources, they would share the nectar information with the onlookers by waggle dance. The onlookers select the food sources by using the nectar information provided by the employed bees in certain probability. The selection probability p_i is calculated

step5: The onlookers searching phase: rnd = rand(0,1) is a random number in the range [0, 1]. If rnd < p_i, the attracted onlookers search in the neighborhood of food sources by using Eq. $v_{i,j} = x_{i,j} + \varphi_{i,j}(x_{i,j} - x_{k,j})$. Otherwise, they search widely according to Eq. (54.8), and memorize the better solution and update vector $Bas(i)$

step6: The scout searching phase: To check whether the searching initial vector $Bas(i)$ reaches the maximum limit search times *limit*, if $Bas(i) > limit$, the *i*th employed bee abandons the current food source and becomes a scout to continue searching globally by using Eq. (54.9)

step7: If $y_{\min} < \min$, set $\min = y_{\min}, X_g = X_g^{'}(t)$. Memorize the best solution x^* as the global optimal solution

step8: Update iteration count, $t = t + 1$. If it meets the current number of iterations $t > MCN$, stop searching, and output minimum $f_{\text{opt}}^{-}(X_g)$ and the global optimal position X_g, otherwise return to Step3 and continue.

54.3 Computer Numerical Simulation Results

54.3.1 Test Functions and Experimental Setups

In order to evaluate the performance of the proposed algorithm, we apply it to nine classical benchmark test functions (Ackley, Griewank, Penalized1, QuarticNoise,

Rastrigin, NC_Rastrigin, Rosenbrock, SphereModel, and Schwefel2.22) [13]. These functions are shorthand for *AC*, *GR*, P_1, *QN*, *RA*, *NR*, *RO*, *SM,* and S_{22}, respectively, which include various types of complex issues, such as the unimodal and multimodal functions, regular and irregular functions, separable and nonseparable functions.

In order to verify the effectiveness of the proposed algorithm, the authors carried out a number of computer numerical simulation experiments, including the basic ABC algorithm with other recently proposed swarm optimization algorithms, such as RABC, PS-ABC, BABC, HHSABC, and NABC to compare. Laboratory equipment for the general laptop computer, with CPU of Intel (R) Core ™ 2 Duo CPU T6500 2.10 GHz, 4G memory, and experimental simulation software is MATLAB7.0.

54.3.2 Parameter Settings

The number of colony scale algorithm *m* and the limit parameter used to control the number of maximal exploitations per food source are set as 100 for all runs. The initial number of employed bees and onlookers are set as 50, respectively. For comparing with other bee colony optimization algorithms fairly, the programs are tested under the same experimental conditions for setting the same maximum iteration count *MCN* and the same maximum number of evaluations. High-dimensional function is tested with dimensions for 30 and 100, respectively, which are convenient for comparing with. Each of the experiments is repeated 30 times starting from a random population distribution with different random seeds. The test results are shown in Tables 54.1, 54.2, 54.3.

54.3.3 Experimental Results

54.3.3.1 Comparing LFABC with Other ABC Algorithms

From Tables 54.1, 54.2, 54.3 and Fig. 54.1, it could be seen that the LFABC1 algorithm with Lévy flights disturbance strategy is outstanding. Theoretical optimal values conducted on 21 functions have been found by the LFABC1 algorithm. And the LFABC algorithm has reached theoretical optimal values on four functions. Overall, LFABC has better performances than other ABC algorithms on most functions. However, LFABC is superexcellent for both high-dimensional functions, when extremum is 0.

54.3.3.2 Robustness Analysis

In the experiment, the maximum number of evaluations is set as 300,000 to verify the right result. Error permitted range is defined by Eq. (54.12). We consider a trial to be successful if the following inequality holds:

Table 54.1 Compare ABC with LFABC1 and LFABC algorithms in the same number of evaluations 300000(D=30)

Sy	ABC				LFABC				
	C.I.	NFE	SR	Time(s)	C.I.	NFE	SR	Time(s)	AR
AC	1119	113059	1	58.907	119	17262	1	1.79	6.55
GR	778	78578	1	48.686	151	32034	1	4.2688	2.45
P_1	545	55025	1	44.024	167	52709	1	7.4256	1.04
QN	–	–	0	182.082	–	–	0	24.765	–
RA	847	85547	1	36.428	85	12340	1	1.0846	6.93
NR	–	–	0	239.315	89	14188	1	1.4366	–
RO	–	–	0	83.132	–	–	0	23.9502	–
SM	653	65993	1	28.646	103	22439	1	1.9844	2.94
S_{22}	1027	103707	1	70.632	95	19197	1	2.079	5.40
AC	–	–	0	331.541	502	75219	1	8.1106	–
GR	2460	248420	1	321.638	565	112266	1	15.9444	2.21
P_1	1892	191112	1	322.370	657	201197	1	58.264	0.95
QN	447	285183	0.2	380.661	–	–	0	29.2656	–
RA	572	297788	0.2	204.769	406	62429	1	5.4008	4.77
NR	–	–	0	6196.885	401	63277	1	5.8532	–
RO	–	–	0	116.639	–	–	0	27.6856	–
SM	2341	236461	1	60.372	448	92220	1	6.1184	2.56
S_{22}	–	–	0	2129.493	406	75426	1	6.5022	–

Table 54.2 Compare LFABC with LFABC1 and RABC in terms of Mean and SD (D=30)

Sy	NFE	RABC		LFABC1		LFABC	
		Mean	SD	Mean	SD	Mean	SD
SP	$1.5*10^5$	9.11E-61	2.11E-60	7.33E-60	1.29E-59	*0*	*0*
S_{22}	$2.0*10^5$	3.21E-74	1.97E-73	5.69E-72	1.27E-71	*0*	*0*
RO	$1.5*10^5$	*8.05E-23*	*3.74E-21*	2.41E+01	1.96E+00	2.37E+01	1.56E+00
ST	$1.5*10^5$	*0*	*0*	*0*	*0*	*0*	*0*
QU	$3.0*10^5$	3.58E-02	6.84E-03	5.36E-06	3.59E-06	*2.36E-06*	*2.78E-06*
RA	$5.0*10^5$	*0*	*0*	*0*	*0*	*0*	*0*
AC	$1.5*10^5$	3.82E-14	4.37E-15	*0*	*0*	*0*	*0*
GR	$2.0*10^5$	2.02E-04	1.43E-03	*0*	*0*	*0*	*0*
P_1	$1.5*10^5$	*1.57E-32*	1.66E-47	6.29E-14	7.19E-14	*1.57E-32*	*0*

$$\left|v^* - v^\alpha\right| < \varepsilon_1 \left|v^\alpha\right| + \varepsilon_2, \tag{54.12}$$

where v^* is the best function value obtained by the algorithm, v^α is the analytical global minimum, and the error is expressed as $\left|v^* - v^\alpha\right|$; the accuracy controlling parameters are set as $\varepsilon_1 = 10^{-4}$ and $\varepsilon_2 = 10^{-6}$ [16] in the permitted error range.

Table 54.3 Compare LFABC with other ABC algorithms ($D=30$)

Sy	PS-ABC		BABC		HHSABC	
	Mean	SD	Mean	SD	Mean	SD
AC	8.88E-16	*0*	1.26E-13	3.48E-14	2.07E-14	3.89E-15
GR	*0*	*0*	4.23E-11	2.16E-11	2.04E-16	4.39E-17
P_1	5.53E-16	8.69E-17	2.85E-30	2.19E-30	2.99E-17	1.13E-18
QN	2.15E-02	6.88E-03	3.20E-02	6.03E-03	1.98E-04	6.40E-05
RA	*0*	*0*	*0*	*0*	*0*	*0*
NR	–	–	*0*	*0*	–	–
RO	*1.5922*	4.4066	–	–	2.34E+00	4.71E+00
RO	–	–	*9.06E-06*	*1.41E-05*	–	–
SM	0	0	1.57E-27	1.14E-27	4.36E-18	9.86E-19
S_{22}	*0*	*0*	3.45E-15	8.79E-16	*0*	*0*
Sy	*NABC*		*LFABC1*		*LFABC*	
	Mean	*SD*	*Mean*	*Mean*	*SD*	*Mean*
AC	3.97E-14	5.12E-15	*0*	*0*	*0*	*0*
GR	1.13E-16	3.39E-16	*0*	*0*	*0*	*0*
P_1	–	–	6.71E-10	1.14E-09	*1.57E-32*	*0*
QN	1.56E-02	3.24E-02	*8.33E-06*	*5.09E-06*	9.47E-06	5.56E-06
RA	*0*	*0*	*0*	*0*	*0*	*0*
NR	–	–	*0*	*0*	*0*	*0*
RO	4.50E-02	2.38E-02	2.43E+01	1.87E+00	2.38E+01	*1.77E+00*
RO	–	–	8.04E-02	1.44E-01	6.38E-03	1.27E-02
SM	4.75E-16	3.86E-16	1.79E-44	3.03E-44	*0*	*0*
S_{22}	1.79E-15	2.53E-15	3.70E-42	4.90E-42	*0*	*0*

For testing the convergence speed, we compare the proposed algorithm with ABC algorithm by using acceleration rate (AR) [16] which is calculated by the Eq. (54.13) as follows:

$$AR = \frac{NFE_{\text{ABC}}}{NFE_{\text{LFABC}}}, \tag{54.13}$$

where NFE_{ABC} and NFE_{LFABC} are the number of function evaluations of the ABC and LFABC algorithms, respectively.

Eighteen different dimensional test problems are used to compare between LFABC and ABC algorithms. Each of the algorithms run 30 times on each of these functions under the maximum number of 300,000 evaluations and the average number of evaluations (NFE), the average convergence iterations (C.I.), the average execution time (Time), Symbol(Sy), D(Dimension), AR and optimizing the success rate (SR, optimization success rate = number of experiments to achieve precision/total number of experiments) of the final results are presented in Table 54.1. The C.I., NFE, SR, Time, and AR are used to compare the efficiency between LFABC1, LFABC, and ABC. The results show that the LFABC algorithm has strong search

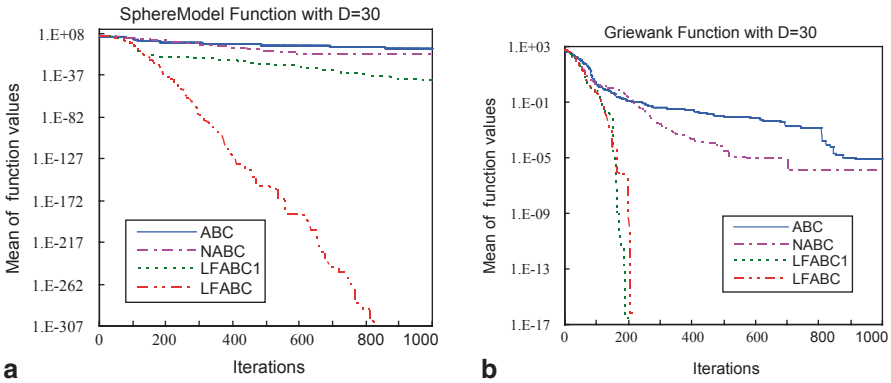


Fig. 54.1 Convergence performance of different ABCs on the two test functions

ability and fast convergence speed for high-dimensional function optimization problems.

54.3.3.3 Performance Comparison Between LFABC and Other Artificial Bee Colony Algorithms

This chapter compares the LFABC algorithm with the latest proposed RABC algorithm, and the experimental parameters are the same with those used in RABC algorithm. As we can see from Table 54.2, the LFABC algorithm is roughly equivalent to the RABC algorithm for low-dimensional functions. However, the LFABC has better performances than the RABC for optimization problems of high-dimensional functions. Its solution precision is far better than the RABC algorithm in most cases except for the Rosenbrock (*RO*) function.

In order to further verify the advantage of the LFABC algorithm, the chapter compares the LFABC algorithm with the latest proposed algorithms such as PS-ABC, BABC, HHSABC, and NABC. From the test results displayed in Table 54.3, we can see that the LFABC algorithm has better performances and higher solution precisions than other existing algorithms tested by most functions.

54.4 Conclusion

A new artificial bee colony algorithm named LFABC based on Lévy flights disturbance strategy is proposed in this chapter. Meanwhile, some modifications to the search mechanism of onlookers and scouts are introduced. In order to evaluate the performances of the proposed algorithm, we apply it to nine classical benchmark test functions. The test results show that the LFABC algorithm has better performances than the ABC algorithm in global search ability, convergence speed, and

solution precision. Meanwhile, with steady optimization performance and strong robustness, the LFABC algorithm is well suitable for solving optimization problems of multimodal and high dimensional functions. It is also an ideal method to solve continuous optimization problems. We believe that the study of LFABC algorithm provides a good theoretical reference value for further research, application and popularization of the artificial bee colony algorithm.

Acknowledgments This chapter is supported by the grants from the National High Technology Research and Development Program of China [Grant No. 2007AA01Z334], the National Natural Science Foundation of China [Grants No. 61272219, 61100110 and 61021062], the Science and technology program of Jiangsu Province [Grants No. BE2010072, BE2011058, BY2012190] and the Innovation Foundation of Huaian College of Information Technology (Grant No. hxyc2013001)

References

1. Karaboga D. An idea based on honey bee swarm for numerical optimization[R], Tech. Rep. TR06, Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.
2. Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm[J]. J Glob Optim. 2007;39(3):459–71.
3. Karaboga D, Akay B. A survey: algorithms simulating bee swarm intelligence[J]. Artif Intell Rev. 2009;31(1–4):61–85.
4. Zhu G, Kwong S. Gbest-guided artificial bee colony algorithm for numerical function optimization[J]. Appl Math Comput. 2010;217(7):3166–73.
5. Achalakul BT, Sirinaovakul B. The best so-far selection in artificial bee colony algorithm[J]. Appl Soft Comput J. 2011;11(2):2888–901.
6. Li B, Li Y, Gong L. Protein secondary structure optimization using an improved artificial bee colony algorithm based on AB off-lattice model[J]. Eng Appl Artif Intell. 2014;27(1):70–9.
7. Gao WF, Liu SY. A modified artificial bee colony algorithm[J]. Comput Oper Res. 2012;39(3):687–97.
8. Karaboga AD. A modified artificial bee colony algorithm for real-parameter optimization[J]. Inf Sci. 2012;192(1):120–42.
9. Li G, Niu P, Xiao X. Development and investigation of efficient artificial bee colony algorithm for numerical function optimization[J]. Appl Soft Comput. 2012;12(1):320–32.
10. Xu YF, Fan P, Yuan L. A simple and efficient artificial bee colony algorithm[J]. Math Probl Eng. 2013;2013(2):1–9.
11. Sharma TK, Pant M, Deep A. Modified foraging process of onlooker bees in artificial bee colony[C]. Proceedings of Seventh International Conference on Bio-Inspired, Computing: Theories and Applications (BIC-TA 2012), Advances in Intelligent Systems and Computing, Springer, Berlin Heidelberg; 2013. pp. 479–487.
12. Gao W, Liu S. Improved artificial bee colony algorithm for global optimization[J]. Inf Process Lett. 2011;111(17):871–82.
13. Kang F, Li J, Ma Z. Rosenbrock artificial bee colony algorithm for accurate global optimization of numerical functions[J]. Inf Sci. 2011;181(16):3508–31.
14. Wu B, Qian C, Ni W, Fan S. Hybrid harmony search and artificial bee colony algorithm for global optimization problems[J]. Comput Math Appl. 2012;64(8):2621–34.
15. Yang XS, Deb S. Cuckoo search via Lévy Flights[C]. Proceedings of World Congress on Nature & Biologically Inspired Computing (NaBic 2009), IIEEE Publications, Coimbatore, India; 2009. pp. 210–214.
16. Kang F, Li JJ, Li HJ. Artificial bee colony algorithm and pattern search hybridized for global optimization[J]. Appl Soft Comput. 2013;13(4):1781–91.

Chapter 55
Improvement of the Positioning Accuracy in GNSS Positioning Based on Neural Network

Zhifei Yang, Yixuan Wang and Ouyang Xu

Abstract Several common technologies in global positioning system (GPS) positioning are compared first, and then a model of dynamic fuzzy neural networks is used in this chapter. It is shown that the dynamic neural network can eliminate residuals and abnormal data in the postprocessing of GPS data effectively; at the same time, the dynamic neural network can also improve the reliability and stability in data processing and thus raise the degree of positioning accuracy.

Keywords Data postprocessing · Dynamic neural network · Relative localization

55.1 Introduction

With the wide application of the global navigation system, more and more electronic devices can obtain their locations by means of satellite positioning, which has greatly facilitated the people's life; however, because of low accuracy of pseudo range positioning and SA (selective availability) policy of the USA, the accuracy of single point positioning could not meet people's demand; therefore, it is important to raise the degree of positioning accuracy. The global positioning system (GPS) relative positioning [1], based on a known precise location called the base station area, can measure the monitor location and get the coordinate components of east, north, highly relative to the base station. To get higher positioning accuracy, it is important to use the pseudo range of the base station and carrier phase information to calculate the entire cycle ambiguity. So far, the methods of least squares and the Kalman filter have been adopted to finish the GPS relative positioning commonly.

Z. Yang (✉) · Y. Wang
Lanzhou Jiaotong University, 730070 Lanzhou, China
e-mail: yzf@mail.lzjtu.cn

Y. Wang
e-mail: roy17007@126.com

O. Xu
State Grid GanSu Maintenance Company, 730050 Lanzhou, China
e-mail: ouyangsanglang@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_55

In the results, there are several abnormal data which may affect the accuracy of the positioning. In this chapter, abnormal points are processed by a dynamic neural network. Empirical results show that the results are smoother and may increase the degree of accuracy with this technology.

55.2 GPS Relative Localization

Relative positioning is known as differential positioning because it uses two or more devices, one called the base station and others called the monitor stations, to receive the common satellite signal synchronously. The position of the base station is unchanged; the relative positions of monitor stations can be obtained by sending the data of the base station to the monitor stations. The base station position can be obtained by averaging values of several measurements as long as the base station position is known; the monitor station's position can be calculated by observing the synchronous satellites. As each receiver observes a common satellite signal synchronously, the satellite clock deviation, ephemeris deviation, and atmospheric delay are the same, which can greatly reduce or even eliminate the system deviation and get a higher positioning precision. The method of carrier phase measurement is usually used in the relative positioning, which can get the accuracy of 5 mm and has been widely used in geodesy, precision engineering measurement, and so on.

There are three kinds of GPS relative positioning methods [2]:

1. The differential of the position

This method calculates the difference of the coordinates on the base station positioned by the GPS receiver and the real coordinates, then gets the differential revised coordinates, which need the same satellite signal between the base station and the monitoring stations, and thus the positioning precision can reach 5–10 m

2. The differential of the pseudo range

It calculates real coordinates of the base station and the satellite constellation to get the pseudo range values; then it compares the values and the pseudo range values from the monitor stations to get D-values as the differential revised coordinates. Because the method revises each satellite's pseudo range, if the signal between the base station and monitor stations is not exactly the common, it can eliminate the differences, thus, the station can capture four or more satellites and the monitor stations can be positioned validly. The accuracy of positioning depends on the number and the distribution of satellites, and the accuracy of this positioning can reach 3–10 m

3. The differential of the carrier phase

The data of carrier phase observations of the base station are passed to the monitoring station through the communication layer; then the monitor station do the positioning of the carrier phase; the accuracy of this positioning can reach the centimeter level.

55.3 Building of Dynamic Fuzzy Neural Networks

55.3.1 Abnormal Data Analyzing

As the bias of the abnormal data is usually larger than the common data, regression of the neural networks can distinguish the differences. As it can remove the abnormal data based on the value of the regression using neural networks, it improves the accuracy of the data.

55.3.2 Structure of Dynamic Fuzzy Neural Networks

Dynamic fuzzy neural networks (D-FNNs) have the ability of parameter adjustment and structure recognition which combines both the advantage of the fuzzy system and the neural networks. Its learning speed is very fast. D-FNN includes four parts: input layer, membership function layer, T-norm layer, and output layer [5] as shown in Fig. 55.1.

1. Input layer

Each node represents an input variable.

2. Membership function layer

Each node-x_i corresponds to u membership function-$A_{ij}(j = 1, 2, 3, \ldots, u)$. The membership function is expressed by Eq. 55.1:

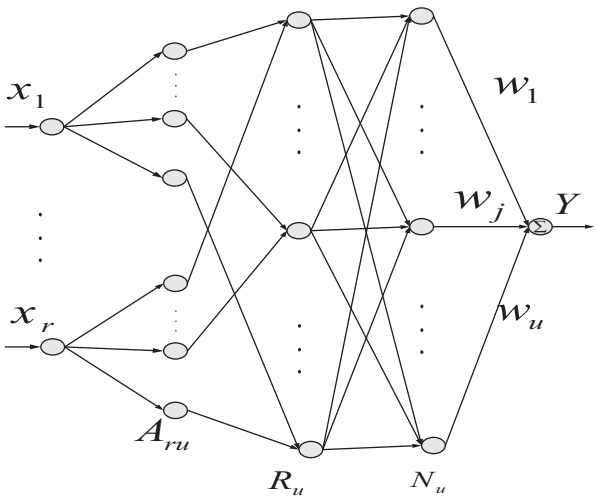


Fig. 55.1 D-FNN structure

$$\mu_{ij}(x_i) = \exp\left[-\frac{(x_i - d_{ij})^2}{\sigma_j^{\,2}}\right](i = 1,2,3,...,r.j = 1,2,3,...,u), \tag{55.1}$$

where μ_{ij} is the j th membership function of x_i, d_{ij} is the j th Gaussian membership function center of x_i, σ_j is the j th Gaussian membership function width of x_i, r is the input variable number, u is the membership function number which represents the total rule numbers of the system.

3. T-norm layer

This layer is responsible for calculating the T-norm operator of every rule triggered weight. The node numbers reflect the fuzzy rule numbers. The j th output of rule R_j is given in Eq. 55.2:

$$\phi_j = \exp\left[-\frac{\sum_{i=1}^{r}(x_i - d_{ij})^2}{\sigma_j^{\,2}}\right](j = 1,2,3,\ldots,u) \tag{55.2}$$

4. Output layer

Every output variable is the sum of all input variables of the system in this layer.

$$y(x_1, x_2, \ldots, x_r) = \sum_{l=1}^{k} w_l \cdot \varphi_l, \tag{55.3}$$

where y is the output of the variable and w_l is the k th rule connecting weight.

55.3.3 Studying the Algorithm of the D-FNN

55.3.3.1 Fuzzy Rule Generation

The fuzzy system has special characteristics. When there are a fewer number of rules, it cannot include the total data information exactly, which will lead to a poor property of the neural networks. If there are too many rules, it will influence the later neural networks' generalization ability, and we would not be able to study the law of the system well [3, 4].

Assume that $X_i = (x_1, x_2, \ldots, x_r)$ is an input vector of the system, t_i is the desired output of the system, and y_i is the output value of the system. Comparing $\varepsilon_i = \|t_i - y_i\|$ and k_e(preselected the desired accuracy), we may determine whether to add a new rule or not [5].

55.3.3.2 Determination of Parameters

The determination of parameters includes the width and center of the Gaussian function. The width of the Gaussian function is exacted by Eq. 55.4:

$$\sigma_j = \frac{\max\{|d_j - d_{j-1}|, |d_j - d_{j+1}|\}}{\sqrt{\ln\left(\frac{1}{\varepsilon}\right)}} \ (j = 1, 2, 3, \ldots, u), \tag{55.4}$$

where $d_j - d_{j-1}$ and $d_j - d_{j+1}$ are the center of two membership function near the j th membership function, ε is the complete parameter of the fuzzy rule.

The center of the Gaussian function is decided via Eq. 55.5:

$$d_{j(u+1)} = x_j^k. \tag{55.5}$$

55.3.3.3 Selecting and Training of the Sample

This chapter selects 30 data as the training sample of the D-FNN system with 7 data for forecasting the comparison from the latest monitoring data. When the model is trained, the steepest descent method is used to modify the weighting coefficients of the network. We can get the shapes of the membership function and network weighting coefficients after adjusting network weights and other adjustable parameters so as to achieve the goal of learning fuzzy rules consequently. The weights-ω are trained as Eq. 55.6:

$$J = \frac{[t_i - y(X_i)]^2}{2} \ (i = 1, 2, 3, \ldots, r). \tag{55.6}$$

Performance indicator- J searches and adjusts the weights in the opposite direction, as shown in Eq. 55.7:

$$\omega_j(k+1) = \omega_j(k) + \alpha\Delta\omega_j(k) - \eta(k)\frac{\partial J}{\partial \omega_j}, \tag{55.7}$$

where α is the inertia coefficient, η is the learning efficiency [6, 7].

The same way may be used to adjust the membership function parameters, such as the shape factor- a_{ik}, as shown in Eq. 55.8:

$$a_{ik}(k+1) = a_{ik}(k) + \alpha\Delta a_{ik}(k) - \eta(k)\frac{\partial J}{\partial a_{ik}}. \tag{55.8}$$

The objective function will achieve optimal expectation after adjusting continuously.

55.4 Relative Localization Combined with the Dynamic Neural Network and Data Analysis

The test used one device as the base station and one device as the monitor station. The base station was on the roof of the library of Lanzhou Jiaotong University. The monitor station was located in a hillside with the distance of about 6 km to the base station. The test was performed for 372 h, and 372 sets of data were accumulated, one set an hour; GAMIT analysis software has been used to get the position results. Figure 55.2 shows the results obtained from GAMIT. In Fig. 55.2, some of the observed data shake seriously, which makes the accuracy lower. In order to correct these abnormal data, the method of the dynamic neural network is adopted; the former 100 data points as training samples, the other 272 points as testing ones. The accuracy is better than before. The result is shown in Fig. 55.3.

As shown in Fig. 55.3, if the abnormal data is dealt correctly, the deformation trend of the data will not be affected.

From Figs. 55.2 and 55.3, it can be known that the dynamic neural network is an effective method to deal with the abnormal data, and data after using this method is closer to the truth value.

Accuracy analysis formula:

$$s = \sqrt{\frac{1}{n}[(x_1 - m)^2 + (x_2 - m)^2 + \ldots.. + (x_n - m)^2]}, \tag{55.9}$$

where m is the average value of $x_1, x_2, \ldots.. x_n$.

The precision before the process of the dynamic neural network is 0.042221.

The precision after the process of the dynamic neural network is 0.038472.

From the above-said precision, it can be known that the degree of accuracy can be raised upon the application of the dynamic neural network.

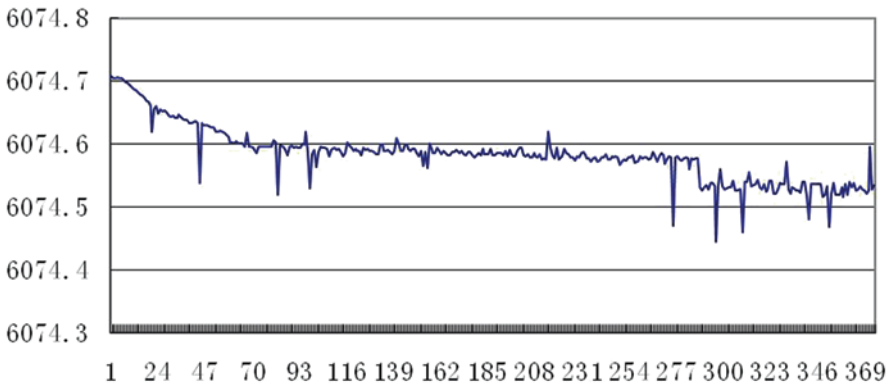


Fig. 55.2 Before the process of the dynamic neural network

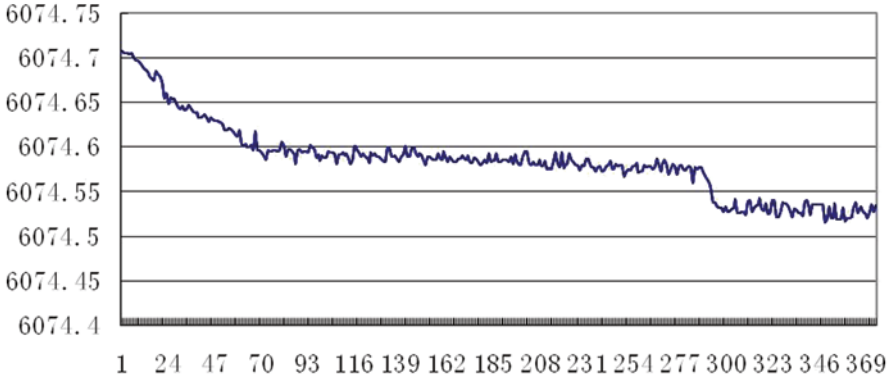


Fig. 55.3 After the process of the dynamic neural network

55.5 Conclusion

The stability of the GPS data can be improved in combination with the dynamic neural network while the abnormal data can be removed and the whole accuracy of data process can be improved. The application of the GPS data combined with the dynamic neural network can satisfy the needs of engineering; however, the current dynamic neural network is only made in the data post process, and it demands further verification to confirm whether it is effective to the real-time kinematic (RTK).

References

1. Zhongmo Z. GPS satellite measuring principle and application. Beijing: Surveying and Mapping Press; 1997. p. 5–200.
2. Dongchen E. High-precision GPS data processing based on Gamit/Globk Software. Polar Rearch. 2005;17(2):173–9.
3. Yang Y, Zhao X, Wang W, et al. Color image filtering method based on fuzzy neural network. Control Decis. 2004;19(1):69–72.
4. Ishibuchi H, Nii M. Numerical analysis of the learning of fuzzified neural networks from fuzzy if then rules. Fuzzy Sets Syst. 2001;120(2):281–307.
5. Haijiao Y, Jun S, Zhao L. The generalized dynamic fuzzy neural networks for stock market prediction. Comput Appl Softw. 2010;27(8):228–31.
6. Er YMJ, Wu S. A fast learning algorithm for parsimonious fuzzy neural systems. Fuzzy Sets Syst. 2002;160(3):337–51.
7. Agiza HN, Yassen MT. Synchronization of Rossler and Chen chaotic dynamical systems using active control. Phys Lett A. 2001;278(4):191–7.

Chapter 56
Fingerprint Indoor Localization Algorithm Based on Modified AdaBoost

Feng Yu, Minghua Jiang, Jing Liang, Xiao Qin, Ming Hu, Tao Peng and Xinrong Hu

Abstract Indoor localization based on fingerprints is becoming increasingly prevalent. In this chapter, a modified AdaBoost algorithm is proposed for localization with received signal strength (RSS). The accuracy of AdaBoost depends on the weak hypothesis from all the weak learning; if there is noise in the fingerprint map, the performance of traditional AdaBoost will decline, and the modified AdaBoost is proposed to remove the error points to enhance the localization accuracy. Experimental results indicate that the modified AdaBoost algorithm obtains high localization accuracy and stability.

Keywords Indoor localization · WiFi · Fingerprint map · Received Signal Strength (RSS) · Modified AdaBoost

M. Jiang (✉) · F. Yu · J. Liang · M. Hu · T. Peng · X. Hu
School of Electronic and Electrical Engineering, Wuhan Textile University, 430200 Wuhan, China
e-mail: yffly@163.com

M. Jiang (✉)
e-mail: jmh@wtu.edu.cn

J. Liang
e-mail: shinelight@wtu.edu.cn

X. Qin
Department of Computer Science and Software Engineering, Auburn University, Auburn, AL, 36849, USA
e-mail: xqin@eng.auburn.edu

M. Hu
e-mail: mhu@wtu.edu.cn

T. Peng
e-mail: pt@wtu.edu.cn

X. Hu
e-mail: hxr@wtu.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_56

56.1 Introduction

Recent advancements in mobile device technology and the increasing availability of WiFi networks have enabled more accurate localization in location-based services (LBSs) such as indoor localization, tracking, navigation, and localization-based security. The global positioning system (GPS) works well in outdoor localization. Unfortunately, GPS satellites are too weak to pass through most constructions, thus making GPS invalid for indoor localization. Many indoor localization algorithms have been proposed, and there has long been interest in the ability to determine the localization of the mobile devices given only the WiFi signal strength. Nowadays, typical indoor localization algorithms based on WiFi networks include time difference of arrival (TDOA) [1], time of arrival (TOA) [2], angle of arrival (AOA) [3], and received signal strength (RSS) [4, 5, 6]. Due to its wide range of application and low cost without additional hardware, the RSS-based algorithm is the best choice for indoor localization.

Currently, the most fashionable RSS-based localization estimation technique is the fingerprint map. A commonly used way of estimating mobile devices is to find the nearest reference point in the fingerprint map algorithm, and using the Euclidean distance in signal space as criteria evaluates the nearest reference point.

Many researchers try to find the best method to enhance localization accuracy and stability. The nearest neighbor (NN) is the simplest nearest neighbor algorithm. For developing localization accuracy, the *k*-nearest neighbor (KNN) is proposed by selecting the nearest *k* neighbors around a device to determine its own position. The fixed number (k) of nearest neighbors may enlarge localization error in the KNN algorithm; sometimes, neighbors far from the mobile devices may be matched in the *k*-nearest neighbor algorithm. Therefore, the weighted *k*-nearest neighbor (WKNN) [7] is proposed by giving each reference point a weight to increase the accuracy of localization. Besides that, the Bayesian probabilistic model (BPM) [8], support vector machine (SVM) [9], multidimensional scaling (MDS) [10], and artificial neural network (ANN) [11] are also proposed in many research studies.

In this chapter, we propose the localization algorithm based on AdaBoost [12, 13, 14] to process the real-time signal strength to match the fingerprint map. The AdaBoost algorithm can enhance the accuracy of the weak learning algorithm, but the accuracy of the classification is not stable when there are noise points in the fingerprint. In order to develop the stability of localization accuracy, the modified AdaBoost algorithm is proposed to remove the error points in the fingerprint map.

The remainder of this chapter is organized as follows: A concise description of indoor localization system is presented in Sect. 56.2; and Sect. 56.3 proposes the modified AdaBoost algorithm. The experimental results and evaluation through implementations are listed in Sect. 56.4. In the end, Sect. 56.5 draws the conclusions.

56.2 Indoor Localization System

As the traditional fingerprint algorithm, the proposed algorithm consists of two different phases: The offline training and the online localization phase. In the first phase, we should build the fingerprint map, and then using the localization algorithm finds the best matching point in the localization phase.

56.2.1 Offline Training Phase

The fingerprint map is built in the offline training phase, and the reference points of RSS are collected at known locations by the mobile devices. Owing to the fact that the same point will detect different RSSs in different orientations, four different orientations (north, south, east, west) are considered in the fingerprint map. The value of RSS measured from *AP i* at *RP j* with orientation *o* is represented as $r_{i,j}^{(o)}(t)$, and *t* represents the total time of measurment. Then, the average value of the RSS is calculated and stored in the fingerprint map.

56.2.2 Online Localization Phase

In this phase, the mobile devices collect the real-time RSS in the WiFi environment, and it is processed by the localization algorithm with the fingerprint map. Since the strong RSS is more stable, we only select the maximum RSS value as localization evaluation standard, determine the orientation of the mobile device by a compass sensor, and match the database using a modified AdaBoost algorithm. According to the modified AdaBoost algorithm, the error points will be moved to develop the localization accuracy, and the algorithm is described in detail in the following section.

56.3 Indoor Localization Based on Modified AdaBoost

56.3.1 AdaBoost Theory

AdaBoost [12] is used to increase the performance of weak learning. It can reduce the deviation of weak learning that consistently produces classifiers which merely need to be a little better than random conjecture. When the training samples $(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)$ are inputting, each x_i represents instance space, y_i represents the classification label that is in the label set Y. We assume $Y = \{-1, +1\}$. A given weak learning trains repeatedly in a series of rounds *t,* for $t = 1, 2, \cdots, T$. $D_t(i)$ represents the weight of distribution on training sample *i* on round *t*. The weight of all the samples is set equally at the beginning. The work of weak learning

is to obtain a weak hypothesis h_t that is evaluated by error ε_t. The detailed introduction is as follows:

- Given training set: $(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)$,
  where $x_i \in X, y_i \in Y = \{-1, +1\}$.
- Initialize $D_1(i) = \frac{1}{m}$.
- For t = 1,2,…, T
- Train weak learning based on distribution D_t
- Obtain weak hypothesis $h_t : X \rightarrow \{-1, +1\}$ with error

$$\varepsilon_t = p_{r_{i\sim D_t}}[h_t(x_i) \neq y_i]. \tag{56.1}$$

- Choose $\alpha_t = \frac{1}{2}\ln(\frac{1-\varepsilon_t}{\varepsilon_t})$.
- Update:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & if \quad h_t(x_i) = y_i \\ e^{\alpha_t} & if \quad h_t(x_i) \neq y_i \end{cases}$$

$$= \frac{D_t(i)\exp(-\alpha_t y_i h_t(X_i)}{Z_t}, \tag{56.2}$$

where Z_t represents the normalization factor.

- Output the ultimate hypothesis:

$$H(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x)). \tag{56.3}$$

56.3.2 Modified AdaBoost Algorithm

The major challenge for precise fingerprint localization comes from the unpredictable and dynamic radio signal, such as diffraction, refraction, and scattering. Under normal circumstances, all the same value RPs are around the testing point; however, sometimes the same value fingerprint points are not all around the testing point. In order to solve this situation, we modify the algorithm based on traditional AdaBoost theory.

Based on the traditional AdaBoost theory in Sect. 3.1, inaccurate noise samples are more difficult to classify. Therefore, along with the iterative increasing, their weight will increase exponentially, and it will reduce the performance of the traditional AdaBoost algorithm. So we remove these error points when the classification accuracy reaches a certain level, and the modified algorithm is as follows (Fig. 56.1).

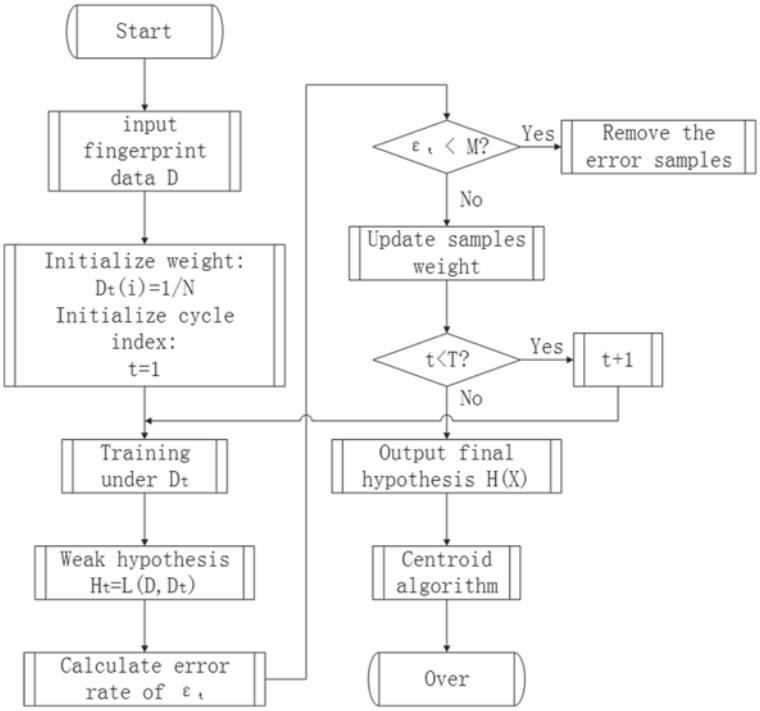


Fig. 56.1 Modified AdaBoost algorithm

56.4 Experimental Results and Evaluation

56.4.1 Experimental Platform Building

We perform the experiment in the third floor of the main building of Wuhan Textile University. The total dimension of the classroom is 10.6 × 6.7 m. The classroom is equipped with a WiFi environment by three TP-LINK TL-MR10U routers. In the experimental environment, we collect 28 points RSS in the centrality of each grid repeatedly and select the mean value.

56.4.2 Performance of the Localization Algorithm

In this work, we compare the proposed modified AdaBoost against the NN, KNN, and WKNN algorithms to estimate the accuracy of the four algorithms in the same data. An arbitrary unknown point is selected as experimental analysis, and the real-time measured RSS is −53 dBm in the actual test. And let $k = 3$, in addition,

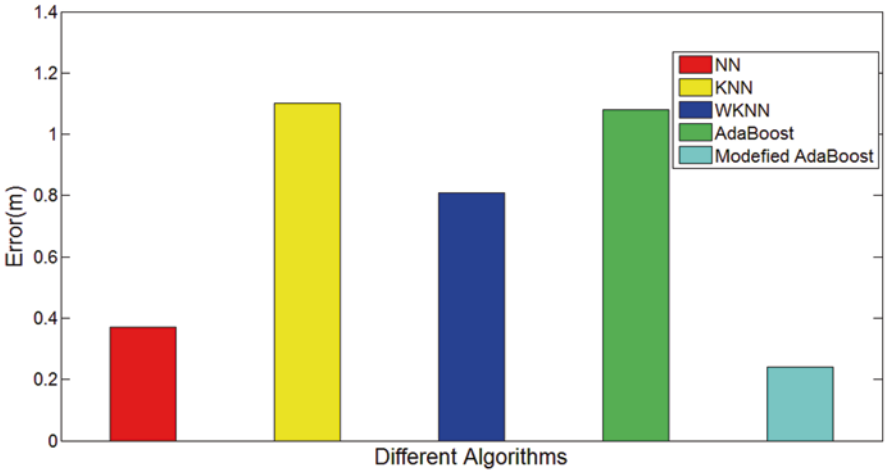


Fig. 56.2 The comparison of localization error

the weighted value $W_i = 1 - 0.1\sigma$, the value of σ is decided by the difference between the measured point and the *k* nearest neighbor points. The localization error is shown in Fig. 56.2.

We get that the modified AdaBoost algorithm is the most accurate of the four; KNN and AdaBoost have the same accuracy, the error of WKNN is a bit smaller than KNN and AdaBoost. Abnormally, NN is more accurate than KNN and WKNN in this experiment. In order to explain the special case, we repeat the experiment 20 times, and select localization accuracy within 2 m as evaluation standard. Figure 56.3 demonstrates the stability of localization accuracy within 2 m.

Figure 56.3 demonstrates that the stability of the five algorithms is that modified AdaBoost > WKNN > AdaBoost > KNN > NN. It can explain that NN is more accurate than KNN and WKNN as shown in Fig. 56.2; in other words, the stability of NN is the worst of all, and we also find that the more close to the routers, the localization accuracy is higher.

From these two figures, we can see that the modified AdaBoost algorithm achieves the highest accuracy of all; not only accuracy, but also the stability of localization accuracy within 2 m. It is powerfully proved that the modified AdaBoost algorithm can improve the localization accuracy.

56.5 Conclusion

In this chapter, a modified AdaBoost algorithm is proposed for indoor localization. The AdaBoost can enhance the accuracy of the weak learning better than random conjecture to a precise strong learning algorithm, bringing about a new method and

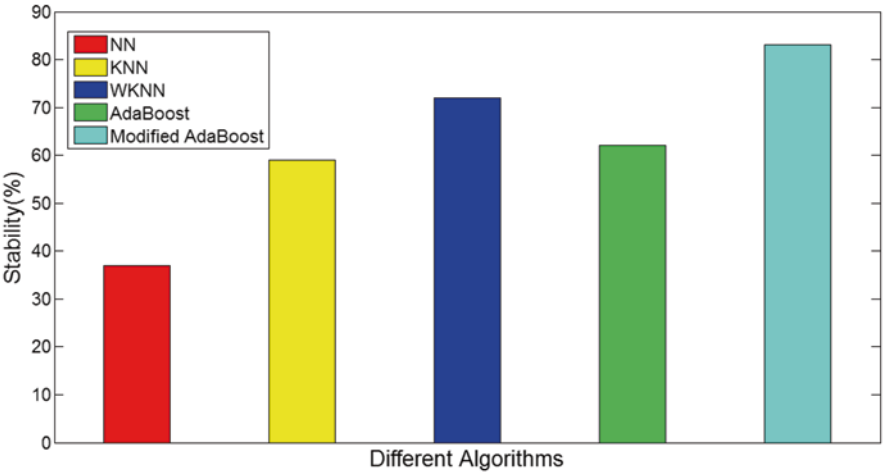


Fig. 56.3 The comparison of localization stability

a new design idea to the design of the learning algorithm. Due to the noise and error, RSS in the fingerprint map can reduce the performance of traditional AdaBoost, and the modified AdaBoost is proposed to solve this trouble by removing these error points, when the classification accuracy reaches a certain level. Experimental results demonstrate that the proposed algorithm can achieve high localization performance. In future, we plan to improve this method to use in more fingerprint databases. Moreover, we intend to test the validity of our proposed method in other complex buildings with many rooms and floors.

Acknowledgments This study was supported by National Natural Science Foundation of China under grant project No. 61232004 and No. 61103085.

References

1. Bocquet M, Loyez C, Benlarbi-Delai A. Using enhanced-TDOA measurement for indoor positioning. Microw Wirel Compon Lett IEEE. 2005;15(10):612–4.
2. Tseng PH, Feng KT, Lin YC, et al. Wireless location tracking algorithms for environments with insufficient signal sources. Mob Comput, IEEE Trans. 2009;8(12):1676–89.
3. Wong C, Klukas R, Messier G. Using WLAN infrastructure for angle-of-arrival indoor user location. Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th; IEEE; 2008. p. 1–5.
4. Serodio C, Coutinho L, Pinto H, et al. A fuzzy logic approach to indoor location using fingerprinting. New York: Springer; 2013. p. 155–69. (Electrical Engineering and Intelligent Systems.)
5. Ouyang RW, Wong AK, Lea CT, et al. Indoor location estimation with reduced calibration exploiting unlabeled data via hybrid generative/discriminative learning. Mob Computing, IEEE Trans. 2012;11(11):1613–26.

6. Yu F, Jiang MH, Liang J, Qin X, et al. An indoor localization of WiFi based on branch-bound algorithm. International Conference on Information Science, Electronics and Electrical Engineering, 2014. ISEEE; IEEE; 2014. 2:1306–8.
7. Shin B, Lee JH, Lee T, et al. Enhanced weighted K-nearest neighbor algorithm for indoor Wi-Fi positioning systems. Computing Technology and Information Management (ICCM), 2012 8th International Conference on; IEEE; 2012. 2:574–7.
8. Madigan D, Einahrawy E, Martin RP, et al. Bayesian indoor positioning systems. INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE; IEEE; 2005. 2:1217–27.
9. Miura H, Sakamoto J, Matsuda N, et al. Adequate RSSI determination method by making use of SVM for indoor localization. Knowledge-Based Intelligent Information and Engineering Systems; Berlin Heidelberg: Springer; 2006. 628–36.
10. Koo J, Cha H. Unsupervised locating of WiFi access points using smartphones. Syst, Man, Cybern, Part C: Appl Rev, IEEE Trans. 2012;42(6):1341–53.
11. Fang SH, Lin TN. Indoor location system based on discriminant-adaptive neural network in IEEE 802.11 environments. Neural Netw, IEEE Trans. 2008;19(11):1973–8.
12. Freund Y, Schapire R, Abe N. A short introduction to boosting. J-Jap Soc Artif Intell. 1999;14(5):771–80.
13. Schapire RE. A brief introduction to boosting. International Joint Conference on Artificial Intelligence; 1999. 99:1401–6.
14. Rätsch G, Onoda T, Müller KR. Soft margins for AdaBoost. Mach Learn. 2001;42(3):287–320.

Chapter 57
Optimization Model of Multitype Generating Units Based on the Latin Hypercube Sampling Method and Mean Lower Semiabsolute Deviation Theory

Haibo Liu, Kangting Chen, Chao Qin, Minfang Huang, Liwei Ju and Zhongfu Tan

Abstract The scheduling optimization of multiple-type generators is of great significance for the economic operation of the power system and the target achievement of energy-saving emission reduction. This chapter builds a two-stage fire and water turbine scheduling the optimization model based on the power performance. Firstly, this chapter simulated the output scenarios of wind power based on the Latin hypercube sampling (LHS) method; then, we minimized the wind curtailment as the main objective, selected the minimization of the system generating energy costs and the generating pollutant emissions as objective functions for cogeneration performance scheduling, and constructed multiple types of generators scheduling the optimization model. The optimization model can guarantee the timely adjustment and output stability for thermal power units in different scenarios, and thus promote the wind power integration.

Keywords LHS · MLSAD (mean lower semiabsolute deviation) · Power generation

57.1 Introduction

In the current market environment of electricity, the power generation scheduling is mainly still in the contract power mode although some provinces have promoted energy-efficient power generation scheduling. Because of the resource constraints, the applicability of energy-efficient power generation scheduling is greatly restricted; thus, we still need to consider the optimization mode of hydrothermal power generation replacement under the contract power mode during the hydrothermal power generation scheduling. As a result, this chapter builds a two-stage fire and

H. Liu (✉) · K. Chen · C. Qin · M. Huang · L. Ju · Z. Tan
School of Economics and Management, North China Electric Power University, 102206 Beijing, China
e-mail: 1196691997@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_57

water turbine scheduling optimization model based on power performance, in the first stage, discusses the relationship between the supply and demand of electricity companies purchase and the national units distribute, sets some different scenarios of generation power supply and demand, and organizes the generation arrangements. Secondly, in order to improve the overall system performance level of power generation, the power trading center brings units together to replace the generation performance of the maximum performance value and gets the actual distribution of power units in the end.

As to the optimization of power generation scheduling, many experts and scholars at home and abroad have experimented a lot of mathematical optimization methods [1], such as the priority method, the dynamic programming method, the intelligent optimization method [2, 3], etc. Literature proposed the programming and decomposition algorithm of annual contract electricity in thermal power units under the three modes of the traditional scheduling, complete market, and limited bidding [4, 5]. In terms of the energy-efficient scheduling, literature reported the study of multiobjective load distribution at the multiload level to solve the configuration optimization problems of distributed generation (DG) in the distribution network, minimizing the network losses and the coal consumption of the generator unit [6]. Literature established a multiobjective generation scheduling optimization model consisting of power system economy, energy saving, and emission reduction and introduced the carbon dioxide emissions into the generation scheduling [7]. The above-said literature studies made an in-depth study of the specific generator unit scheduling model and its cost and benefit, but there is no research on the scheduling optimization model of multitype generating units.

Based on this research background, this chapter uses the Latin hypercube sampling (LHS) method for the random simulation of the wind power output and uses the mean lower semiabsolute deviation (MLSAD) theory for the scheduling optimization of multitype generating units.

57.2 Simulation of the Wind Power Output Scene Based on the LHS Method

The method of Monte Carlo is commonly used in the wind power simulation. As the most classical simulation method of wind power, the Monte Carlo method needs a large number of statistical data and generates random samples to simulate. This requires a lot of efforts to prepare the wind power simulation. In order to solve the above-said problems, LHS is proposed and has been widely used in the power system. This chapter selects the LHS method for stochastic simulation of wind power generation, and the simulation process is as follows:

If the probability distribution function of each element in the K dimensional random variable X is $F_i(i = 1, 2, \ldots, K)$, each element of the vector X is independent of each other, now each element is proceeded to N times the sampling, so we can obtain sampling numerical of each element that is $x_{jk}(k = 1, 2, \ldots, K; j = 1, 2, \ldots, N)$; after

N times of the sampling, we can get a sample matrix of $[P]_{N\times K}$; it supposes the random variable ξ_{jk} is uniformly distributed and the distribution range is [0, 1]; thus, the sampling result is:

$$x_{jk} = F_k^{-1}[(p_{jk} - 1 + \xi_{jk}) / N]. \tag{57.1}$$

In this formula, p_{jk} is the element of j rows and k columns in a matrix P of dimension $N\times K$.

The mean of the function and the function are defined as $E(h(x))$, $h(x)$, and the unbiased estimate of $E(h(x))$ is

$$\hat{h} = \sum_{j=1}^{N} h(x_j) / N. \tag{57.2}$$

So the variance of unbiased estimation $\hat{h}$ when simple random sampling is

$$D(\hat{h}) = D(h(x)) / N. \tag{57.3}$$

The variance of unbiased estimation $\hat{h}$ of LHS is

$$D(\hat{h}) = D(h(x)) / N + (N-1)\operatorname{cov}(h(x_{1,n}), h(x_{2,n})) / N. \tag{57.4}$$

We can see that the number of samples generated by LHS is decreased significantly; but in practice, the number of samples should be compressed further.

57.3 Multiple Types of Generators Scheduling the Optimization Model Based on the MLSAD

57.3.1 MLSAD Model

The joint operation optimization problem of wind power, thermal power, and pumped storage power station can be understood as a unit output combination problem to optimize the proportion of each unit output in total load demand; in this sense, this problem can be solved by a traditional portfolio model. The traditional portfolio model targets at the mean-variance minimization so that the investment decision could achieve the highest revenue effect at the lowest level of risk, but its process of solving is complicated. To simplify the solving process of the portfolio model, Konono and Yamazaki proposed the mean-absolute deviation portfolio model and proved that the model can not only simplify the process of solving but also meet the goal of maximizing the expectation effect. But scholars studying the investment behavior noted that investors would be more concerned about the level of investment risk at the moment of making investment decisions; besides, they

would understand it as a risk when the actual income is lower than the expected revenue. The traditional model portfolios and the mean-absolute deviation portfolio model have equal treatment to the positive and negative deviations from the expected value; consequently, they cannot effectively react to the investor risk appetite. In order to solve this problem, the MLSAD model has been proposed and is widely used in the power system. This model is detailed as below:

Assuming that there are n kinds of assets in a portfolio problem, the investment return rate of each asset is $R_i (i = 1, 2, \ldots, n)$. x_i is the proportion of each asset in the investment portfolio, so the rate of return on investment and expectation are calculated as below in the investment portfolio of $x = (x_1, x_2, \ldots, x_n)^T$:

$$R(x) = \sum_{i=1}^{n} R_i x_i \tag{57.5}$$

$$\mu = E(R(x)) = E\left(\sum_{i=1}^{n} R_i x_i \right) = \sum_{i=1}^{n} E(R_i) x_i. \tag{57.6}$$

57.3.2 Model Construction

57.3.2.1 Objective Function

This chapter takes the minimization of wind curtailment as the main objective and selects the minimization of the system generating energy costs and the generating pollutant emissions as the objective function for cogeneration performance scheduling. The specific objective functions are shown as below:

1. The objective function of minimizing wind curtailment

$$\text{(P1)} \quad \min f_1^s = \sum_{t=1}^{T} \left(\left(g_{wt}^* \right)^s - g_{wt}^s \right). \tag{57.7}$$

In this formula, f_1^s is the objective function of minimizing wind curtailment in scene *s,* $(g_{wt}^*)^s$ is the available power output of wind power in scene *s*.

2. Objective function of minimizing generating energy costs

$$\text{(P2)} \quad \min f_2^s = \sum_{t=1}^{T} \sum_{i=1}^{I} \left(\alpha_i + \beta_i g_{it}^s + \delta_i \left(g_{it}^s \right)^2 \right). \tag{57.8}$$

In this formula, f_2^s is the objective function of minimizing generating costs; g_{it}^s is the generation output of thermal power *i* in scene *s* at time *t; I* is thermal power collection; α_i, β_i, and δ_i are power consumption coefficients, and they are obtained from the historical data regression of generating units.

3. Objective function of minimizing generating emissions

$$\text{(P3)}\ \min f_3^s = \sum_{t=1}^{T}\sum_{i=1}^{I}\left(\omega_i + \xi_i g_{it}^s + \psi_i \left(g_{it}^s\right)^2\right). \tag{57.9}$$

In this formula, f_3^s is the objective function of minimizing generating emissions, ω_i, ξ_i, and ψ_i are generating emission coefficients. They are obtained from the historical data regression of generating units.

57.3.2.2 Constraints

1. System power balance constraint

$$\sum_{i=1}^{I} g_{it}^s + \sum_{w=1}^{W} g_{wt}^s + \sum_{h=1}^{H} g_{ht}^s = L(t). \tag{57.10}$$

In this formula, $L(t)$ indicates the system load demand at time t, g_{ht}^s is the power output of pumped storage generators at time t.

2. Pumped storage unit power output

It is assumed that the relationship between the active and flow of the pumped storage power station is a quadratic function:

$$g_{ht}^s = a_h + b_h q_{ht}^s + \psi_h \left(q_{ht}^s\right)^2. \tag{57.11}$$

In this formula, a_h, b_h, and ψ_h are the power output coefficients of the pumped storage power station.

3. Operating constraints of pumped storage units

$$g_{ht,\min}^s \le g_{ht}^s \le g_{ht,\max}^s \tag{57.12}$$

$$q_{ht,\min}^s \le q_{ht}^s \le q_{ht,\max}^s. \tag{57.13}$$

In this formula, $g_{ht,\min}^s$ and $g_{ht,\max}^s$ are the minimum and maximum power outputs of the pumped storage unit h in scene s at time t, $q_{ht,\min}^s$ and $q_{ht,\max}^s$ are the minimum and maximum power flow (negative) or pumping flow (positive) of the pumped storage unit h in scene s at the time t.

1. Drainage water constraints

$$\sum_{t\in T_h^{\text{gen}}} q_{ht}^s = W_h^{\text{gen}} \le W_{h,\max}^{\text{gen}} \tag{57.14}$$

$$-\sum_{t\in T_h^{\text{pump}}} q_{ht}^s = W_h^{\text{pump}} \geq W_{h,\min}^{\text{pump}} \tag{57.15}$$

$$W_h^{\text{pump}} = W_h^{\text{gen}}. \tag{57.16}$$

In this formula, T_h^{gen} and T_h^{pump} are the period of generating and electricity time collection of the pumped storage unit h. W_h^{gen} and W_h^{pump} are the water consumption and water flowage of the pumped storage unit h in the scheduling period. $W_{h,\max}^{\text{gen}}$ and $W_{h,\min}^{\text{pump}}$ are the maximum water consumption and the minimum water flowage of the pumped storage unit h in the scheduling period.

2. System spinning reserve constraint

$$\sum_{i\in I}(g_{it,\max} - g_{it}^s) + \sum_{h\in H}(g_{ht,\max} - g_{ht}^s) \geq R'(t). \tag{57.17}$$

In this formula, $R'(t)$ is the spinning reserve demand of the system at time t

3. Thermal power adjustment constraints

In order to ensure that the power plant can timely adjust in different scenarios and promote wind power integration, we must consider the difference between the thermal power output and its average output of each scene, so that we can ensure the relative stability of the thermal power output through the constraints of difference. The specific constraints are shown as below:

$$g_{it}^{avg} = \sum_{s=1}^{S} p_s g_{it}^s \tag{57.18}$$

$$-\Delta g \leq g_{it}^s - g_{it}^{avg} \leq \Delta g. \tag{57.19}$$

In these formulas, g_{it}^{avg} is the mean output of thermal power i in each scene. p_s is the probability of occurrence of scene s. $-\Delta g$ and Δg are the upper and lower limits of quick adjustment in thermal power i.

57.4 Solving Process

1. With the models of (P1), (P2), (P3), we can get the generator output and the objective function value in different objective functions so that an objective function input–output table can be available with details shown in Table 57.1

Table 57.1 Input–output table of objective function

Objective Function	f_1	f_2	f_3
f_1^*	$f_1^{\min}$	$f_1^{(2)}$	$f_1^{(3)}$
f_2^*	$f_2^{(1)}$	$f_2^{\min}$	$f_2^{(3)}$
f_3^*	$f_3^{(1)}$	$f_3^{(2)}$	f_3^{min}

2. The main objective function is to minimize the cost of power consumption, the wind curtailment, and the pollutant emissions with specific steps as shown below:

1. Expectation of generating energy cost

$$\overline{f}_2 = \sum_{s=1}^{S} p_s f_2^s \tag{57.20}$$

2. Semiabsolute yield spread of generating energy cost

$$E = \sum_{s=1}^{S} p_s \left| \overline{f}_2 - f_2^s \right| \tag{57.21}$$

3. The critical value of wind curtailment volume and power waste emissions should be ensured. The solution set is satisfactory only when the objective function value is less than the critical value.

$$\overline{f}_1 = \sum_{s=1}^{S} p_s f_1^s \tag{57.22}$$

$$\varpi^s = \min\left\{ f_1^{(2)s}, f_1^{(3)s} \right\} \tag{57.23}$$

$$\varpi = \sum_{s=1}^{S} p_s \varpi^s. \tag{57.24}$$

In these formulas, $\overline{f}_1$ is the objective function expectation of the minimum wind curtailment volume. ϖ^s is the critical value of objective function of the minimum wind curtailment in the scene s. ϖ is the critical value of objective function of the minimum wind curtailment.

$$\overline{f}_3 = \sum_{s=1}^{S} p_s f_3^s \tag{57.25}$$

$$\vartheta^s = \min\left\{ f_3^{(1)s}, f_3^{(2)s} \right\} \tag{57.26}$$

$$\vartheta = \sum_{s=1}^{S} p_s \vartheta^s. \tag{57.27}$$

In these formulas, $\overline{f}_2$ is the objective function expectation of the minimum generating emissions. ϑ^s is the critical value of objective function of the minimum generating emissions in the scene s. ϖ is the critical value of objective function of the minimum generating emissions.

57.5 Conclusion

To optimize the joint operation of the wind power, the thermal power, and the pumped storage power plant, this chapter builds the scheduling optimization model of multitype generating units based on the MLSAD theory. Firstly, this chapter simulates the output scenarios of wind power based on the LHS method, introduces the Kantorovich distance to cut scenes, and removes the small probability scene to replace the original scene with less scene; in this sense, this topic is of practical significance. Considering the wind curtailment, the generating energy costs, and the generating pollutant emissions, this chapter optimizes the operation of multitype generating units so as to ensure that the power plant can be timely adjusted in different scenarios and promote the wind power integration.

Acknowledgments This chapter is supported by the Project of Beijing Municipal Philosophy and Social Science Planning (12JGC080).

References

1. Wang N. Study on optimization model and method of generation dispatching. Beijing: North China Electric Power University; 2011.
2. Simopoulos DN, Kavatza SD, Vournas CD. Unit commitment by an enhanced simulated annealing algorithm. IEEE Trans Power Syst. 2006;21(l):68–76.
3. Simon SP, Padhy NP, Anand RS. An ant colony system approach for unit commitment problem. IEEE Trans Power Syst. 2006;28(5):315–23.
4. Chen G, Xu Y. Research on the operation of cascade power stations based on contract power quantity. China Rural Water Hydropower. 2011;2011(5):159–64. (In Chinese)
5. Zhang L, Liu J, Liu J, et al. Study on scheduling and resolution algorithm of annual contract volume for thermal power units. Relay. 2007;35(4):64–9.
6. Su P, Liu T, Zhao G, et al. An improved particle swarm optimization based multi-objective load dispatch under energy conservation dispatching. Power Syst Technol. 2009;33(5):48–53. (In Chinese)
7. Zhang J, Liu N, Shan L, et al. Research on multi-objective optimization power dispatch based on carbon emissions and its application. J Northeast Dianli Univ. 2013;33(6):6–10. (In Chinese)

Chapter 58
Direction of Arrival Estimation of Two Wide-Band Sources with a Four-Element Planar Cross Array Based on Three Algorithms

Xiaojuan Bai, Yong Shen, Hui Yang, Yansheng Liu, Hongqiang Wang and Xiao Hu

Abstract The acoustic signal of a very low-flying armed helicopter is mainly characterized with a low-frequency and wide-band signal. In this chapter, the beamspace coherent signal-subspace method (BSCSM) is studied with a small size four-element planar cross array. This array is applied to the orientation of two wide-band sources, and its performance is analyzed. The BSCSM, incoherent signal-subspace method (ICSM), and coherent signal-subspace method (CSM) are employed to estimate sources' directions of arrival (DOAs). Simulation results show that BSCSM shows best performance among the three algorithms, and the CSM shows performance which is better than the ICSM does in the acoustic passive direction estimate of two wide-band sources in the far-field environment with a small size four-element planar cross array.

Keywords Acoustic passive direction estimate · Four-element planar cross array · Signal subspace

58.1 Introduction

Research on directions of arrival (DOA) estimation [1] for low-flying armed helicopters becomes an urgent demand in modern war. Radar systems are limited in detection and orientation of low-flying targets in the air. Thus, according to the acoustic signals produced by flying armed helicopters, research and design of a smart sound probe system with automatic detection, identification, orientation, and tracking based on acoustic passive direction finding technology [2] will be essential. Direction finding for a single signal is already studied thoroughly, but in reality there are often multiple targets simultaneously, and the studies are relatively few. Therefore, the study of two targets for the intelligent acoustic probe system is more

X. Bai (✉) · Y. Shen · H. Yang · Y. Liu · H. Wang · X. Hu
95997 Troops of Chinese PLA, 100076 Beijing, China
e-mail: 6266927@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_58

significant. Currently, there are three basic methods in direction finding: the steerable beamforming technology based on the maximum output power, the direction finding technology based on acoustic arrival time difference,and high-resolution orientation estimation techniques. Since the 1970s, high-resolution DOA estimation has become one of the important research directions in the sensor array signal processing [3].

The acoustic signal of a very low-flying armed helicopter is mainly with a low-frequency and wide-band signal. The incoherent signal-subspace method (ICSM) [4] and the coherent signal-subspace method (CSM) [5] of element space are widely adopted for the DOA estimation of two wide-band sources. Among them, the ICSM first breaks down the received broadband data into several narrowband components. Then the position is estimated for each narrowband component. And finally, the result of each narrowband component estimation is combined to achieve the DOA estimation of broadband sources. However, the ICSM is not able to deal with coherent broadband source signals, which can be achieved by the CSM. The principle of the CSM is that it divides coherent broadband data into a number of narrowband components, and then uses the focusing matrix method which focuses various frequency components to the reference frequency. Finally, the narrowband subspace method is used to achieve the DOA estimation. Compared to the ICSM, the CSM will fully extract information which improves detection capability and provides better resolution in spite of low signal-to-noise ratio (SNR). However, both the system errors are too sensitive, and the computation load is relatively large. Beam-space algorithm can reduce the threshold and computation load, and can reduce the system error. In order to obtain DOA estimation of two broadband sources, we select the beam-space coherent signal-subspace method (BSCSM) [6] to process low-flying acoustic targets, which is based on the small-scale four-element planar cross array. Beam pretreatment is used to get source orientation pre-estimates. And during the beam-field orientation estimate, the constant beam-width broadband beam-space high-resolution algorithm [7] is employed to ensure the beam output. The direction estimation and direction finding performance of the BSCSM on the dual-broadband signal source is simulated. The BSCSM, ICSM, and CSM are employed to estimate sources' DOAs. The performance of the BSCSM is compared with the ICSM and CSM using computer simulation. Simulation results show that the BSCSM shows best performance of the three algorithms, and the CSM shows performance better than the ICSM.

58.2 Broadband Array Signal Modeling

The basic principle of acoustic source passive direction finding is the use of acoustic array to receive the information of the target sound, and then to combine with effective algorithms, to achieve spatial orientation of the target, such as target azimuth, pitch angle, etc. We know that the path of the sound waves of the goal is different, and thus the time of each signal received by each element is different, which

forms the time delay. Because the time delay associates with the arrival angle of the target, the delay information is the basis for the passive orientation system of the airborne sound [8].

The four-element planar cross array receives the far-field signal of the two broadband sources of radiation. The coordinate system is shown in Fig. 58.1: The binary linear array A, C and B, D are mutually orthogonal, and the intersection is the coordinate origin. Cartesian coordinates of each element of the four-element planar cross array are A (d,0,0), B(0,d,0), C(−d,0,0), and D(0,−d,0), respectively. The distance between the far-field broadband source M and the receiving array is much larger than the array element spacing 2d, and the source signal received by array can be assumed a plane wave. The spherical coordinates of the sound source is M(r,θ,φ), and P is the projection of M on the XOY plane; so the distance of the target sound source and the coordinate origin is r, the azimuth angle is θ (0° ≤ θ < 360°), and the pitch angle is φ (0° ≤ φ ≤ 90°). We know that the helicopter is warning in the take-off stage; therefore, we assume that the pitch angle is approximately 90°, the azimuth angle on the horizontal XOY will be estimated [9].

Let us assume that a sensor array consists of M sensors receiving plane wavefronts generated by D wideband sources at locations θ_1, θ_2, ..., θ_D, in the presence of additive white noise. The form of frequency domain output can be denoted as matrix form, and its expression can be given by

$$X(f_j) = A(f_j,\Theta)S(f_j) + N(f_j), \tag{58.1}$$

where

$$X(f_j) = [x_1(f_j)x_2(f_j)\ldots x_M(f_j)]^T \tag{58.2}$$

$$N(f_j) = \left[n_1(f_j) \quad n_2(f_j) \quad \cdots \quad n_M(f_j)\right]^T \tag{58.3}$$

$$S(f_j) = \left[s_1(f_j) \quad s_2(f_j) \quad \cdots \quad s_D(f_j)\right]^T. \tag{58.4}$$

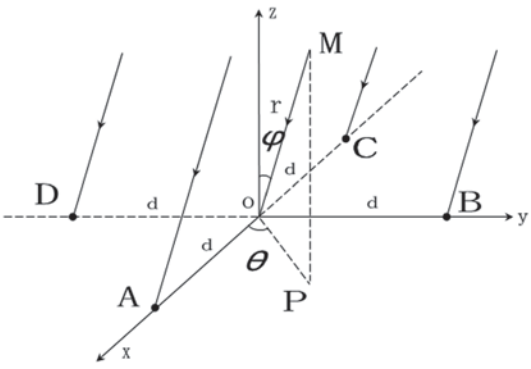


Fig. 58.1 Model of receiving signal from far field with a four-element planar cross

Here X (f_j), N (f_j), and S (f_j) are the output signal vector of the array, the additive white noise vector of the array, and Fourier transformation of the received signal, respectively, at frequency f_j. A (f_j,Θ) = [a(θ1), a(θ2), …, a(θD)] denotes the array manifold vectors, at the frequency f_j and angle Θ. Θ = [θ1, θ2, …, θD] denotes the incidence direction of the D signals.

It is assumed that the signal and noise received by the acoustic array are unrelated, so the output cross-spectral density matrix (CSDM) by array at the frequency f_j can be given as

$$\mathrm{R}(\mathrm{f_j}) = \mathrm{E}\left\{\mathrm{X}(\mathrm{f_j})\mathrm{X}^{\mathrm{H}}(\mathrm{f_j})\right\} = \mathrm{A}(\mathrm{f_j})\mathrm{R_S}(\mathrm{f_j})\mathrm{A}^{\mathrm{H}}(\mathrm{f_j}) + \mathrm{R_n}(\mathrm{f_j}), \tag{58.5}$$

where

$$\mathrm{R_s}(\mathrm{f_j}) = \mathrm{E}\left\{\mathrm{S}(\mathrm{f_j})\mathrm{S}^{\mathrm{H}}(\mathrm{f_j})\right\} \tag{58.6}$$

$$\mathrm{R_n}(\mathrm{f_j}) = \mathrm{E}\left\{\mathrm{N}(\mathrm{f_j})\mathrm{N}^{\mathrm{H}}(\mathrm{f_j})\right\}. \tag{58.7}$$

Here E and H are mathematical expectation and vector or matrix conjugate transpose, respectively. Rs(f_j) and Rn(f_j) are the signal correlation matrix, and the noise correlation matrix, respectively, at the frequency f_j.

58.3 DOA Estimation Based on Constant Beam-Width Beamformers

Here the frequency domain output of the array is divided into J narrow sub-bands in the frequency domain which is used to emulate two broadband source signals received by a four-element planar cross array. The frequency domain outputs of constant beam-width beamformers about the *jth* (j = 1, 2, …, J) narrow sub-band can be given by

$$\mathrm{Y}(\mathrm{f_j}) = \mathrm{W}^{\mathrm{H}}(\mathrm{f_j})\mathrm{X}(\mathrm{f_j}) = \mathrm{W}^{\mathrm{H}}(\mathrm{f_j})\mathrm{A}(\mathrm{f_j})\mathrm{S}(\mathrm{f_j}) + \mathrm{W}^{\mathrm{H}}(\mathrm{f_j})\mathrm{N}(\mathrm{f_j}), \tag{58.8}$$

where

$$\mathrm{W}^{\mathrm{H}}(\mathrm{f_j})\mathrm{N}(\mathrm{f_j}) = \mathrm{N_B}(\mathrm{f_j}) \tag{58.9}$$

$$\mathrm{W}(\mathrm{f_j}) = \left[\mathrm{w_1}(\mathrm{f_j}) \quad \mathrm{w_2}(\mathrm{f_j}) \quad \cdots \quad \mathrm{w_B}(\mathrm{f_j})\right]. \tag{58.10}$$

Formula (10) is the beamforming matrix of the B conjunctive beamformers. Because the beamformers are designed to have constant beam-width over the interested frequency band, W (f_j) satisfies the following relationship:

$$W^H(f_j)A(f_j) \cong W^H(f_0)A(f_0), \tag{58.11}$$

where f_0 is the reference frequency, Eq. (58.11) is a necessary condition of constant beam-width design. The average CSDM is given by

$$\begin{aligned} R_y(f_j) &= E\left\{W^H(f_j)X(f_j)X^H(f_j)W(f_j)\right\} \\ &\cong W^H(f_0)A(f_0)\frac{1}{J}\sum_{j=1}^{J} a_j R_s(f_j)A^H(f_0)W(f_0) \\ &+\frac{1}{J}\sum_{j=1}^{J} a_j W^H(f_j)R_n(f_j)W^H(f_j), \end{aligned} \tag{58.12}$$

where a_j is the scalar, the weighting factor of each narrow sub-band of the CSDM. The matrix pair is given by

$$\{R_Y, \sum_{j=1}^{J} a_j W^H(f_j)R_n(f_j)W^H(f_j)\}. \tag{58.13}$$

The matrix pair can be decomposed for generalized eigenvalue decomposition, and noise subspace $\hat{E}_{Bn}$ is built. A azimuth spectrum can be given as follows:

$$P_{MUSIC}(\theta) = \frac{1}{a^H(f_0,\theta)W(f_0)\hat{E}_{Bn}\hat{E}_{Bn}{}^H W^H(f_0)a(f_0,\theta)}, \tag{58.14}$$

where W^H (f0) a (f0, θ) is the transformed steering vector in the constant beam-width beam-space. The source directions are given by the P peak positions of the MUSIC spatial spectrum.

58.4 Simulation

In the following simulation, we use the array as shown Fig. 58.1. The distance of the array elements A and C equals 1 m, and the distance of array elements B and D equals 1 m, that is d = 0.5 m. The incident signal is the stationary Gaussian random process, and the power of each frequency component is the same.

In order to get the real signal data, we intercept the period of the measured helicopter noise in the testing process during the simulation process. The sampling rate is 10 KHz and the sampling points are 1000 points. Interception in the frequency spectrum is between 0 and 1000 Hz as shown in Fig. 58.2. We can see that the

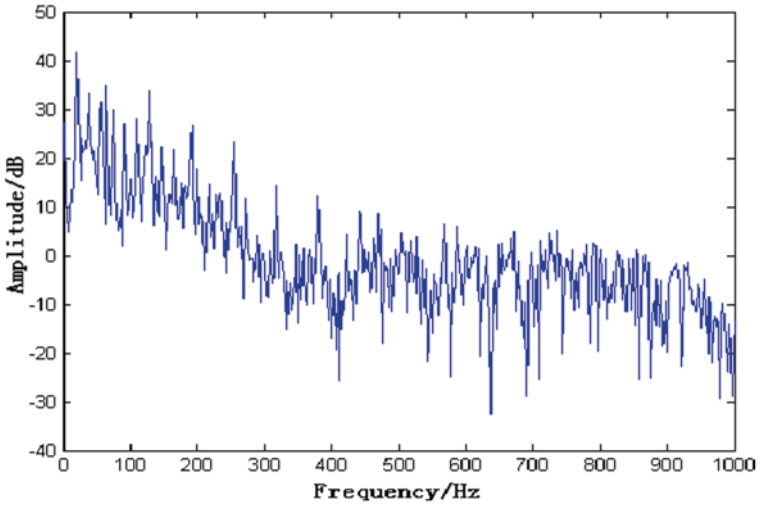


Fig. 58.2 Noise spectrum of the *low*-flying armed helicopters

helicopter noise energy is concentrated in the range of 0−250 Hz. We intercept the helicopter noise as the incoming signal *s(t)*. For each element of the array, the incident signal on each element is generated by simulation, and the white noise *n (t)* in each element is added to the incident signal to get different SNR. Thus, we can get the sampled array output data *x(t)*. The incidence direction θ of the signal can be arbitrarily selected, and SNR can be adjusted by changing the power of additive white noise.

Since the element number of four-element planar cross array is only four and the array aperture is relatively small, the 125−175 Hz broadband signal is divided into 20 narrow sub-bands, where 150 Hz is the reference frequency. Azimuth angles of two broadband signal sources are -5° and 40°. The −3dB beam-width on reference frequency 150 Hz of a four-element planar cross array is 84°, i.e., the width of the main lobe is large, so we select 0°, $\pm 60^{\circ}$, $\pm 120^{\circ}$ in the preformed beam multibeam system. Then we design the beam pattern of five directions of the beam in 20 narrow sub-band constant beam width. Figure 58.3 only points the 00, 600, and 1200 beam patterns; the −600 and −1200 beam patterns are symmetrical and identical with the 600 and 1200 beam patterns. The effect can be seen from the figure, the constant beam-width beam pointing 60°and 120° are not better than the constant beam-width beams of pointing 00, and the side lobe of the constant beam-width beam is relatively high. Figure 58.4 shows the evaluation of two signals' azimuth angle by the BSCSM.

In order to compare the abilities to distinguish the BSCSM, ICSM, and CSM about two broadband sources which are close to each other, let us assume that the azimuth angles of the incident directions of the two broadband signal sources

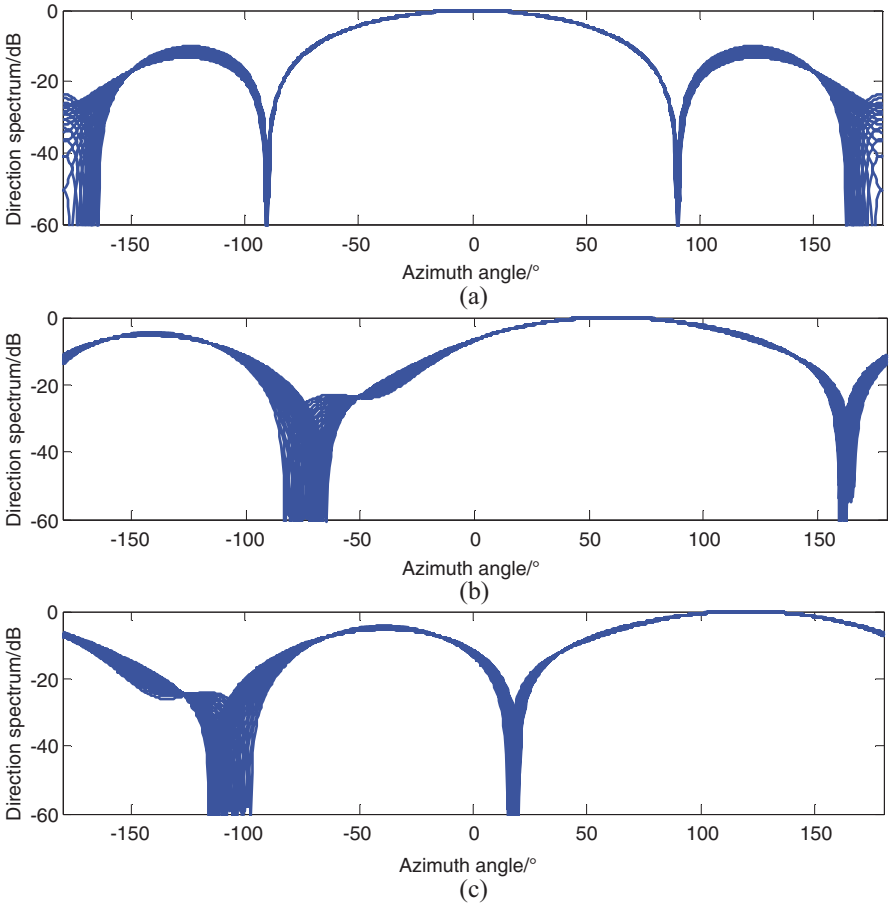


Fig. 58.3 Constant beam-width beam pattern when the direction equals 0°, 60°, 120°

are −50 and 400, respectively, and SNR = 5dB. The broadband signal source is generated as above. The ICSM is divided by the frequency band in the range of 125–175 Hz into 20 narrowband components, and a total of 21 frequency points for DOA estimation are taken. The CSM uses the rotation invariant subspace focusing method which belongs to a nonfocused loss method. The BSCSM is as above [10]. Computation load of the BSCSM is the smallest among the three algorithms in the calculation process.

It can be seen in Fig. 58.5 that the azimuth angles of two broadband sources are estimated by the BSCSM, ICSM, and CSM. The corresponding real curve has two peaks, and they are −5and 40°, i.e., from the figure, we can see that the orientation spectrum background of the BSCSM is significantly lower than the CSM, and the orientation spectrum background of the CSM is significantly lower than the ICSM.

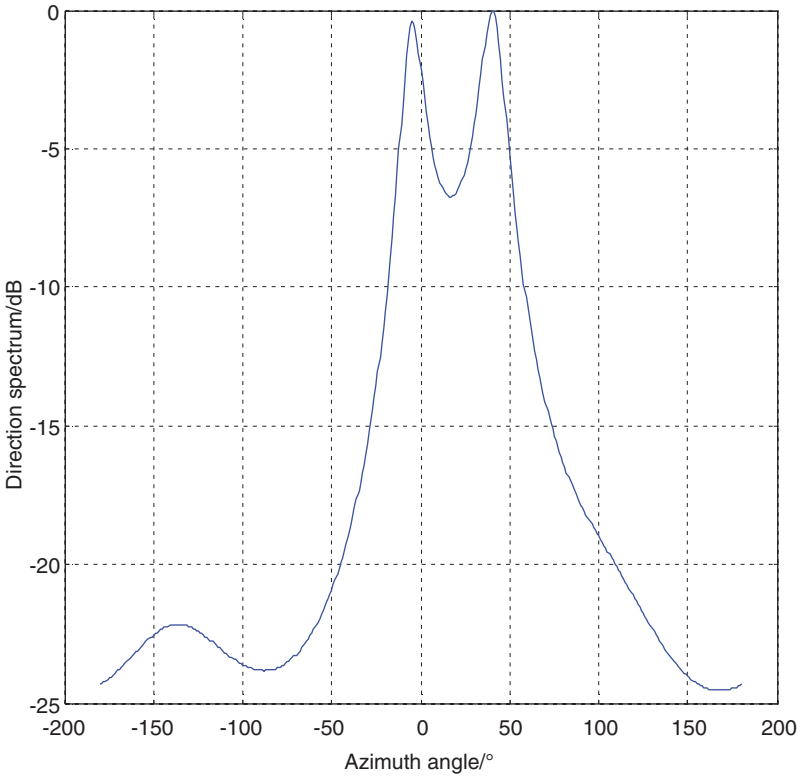


Fig. 58.4 Evaluation of two signals' azimuth angle by the BSCSM

BSCSM's resolution is higher than CSM's, and CSM's resolution is higher than ICSM's.

Figure 58.6 shows the deviation comparison among BSCSM, ICSM, and CSM, and Fig. 58.7 shows the standard deviation estimated by the three methods. Deviation and standard deviation in the figure are estimated through the arithmetic mean of the bias and standard deviation. It can be seen that the BSCSM's deviation is smaller than the CSM in the high SNR, while the standard deviations are considerable. And the CSM's deviation estimates smaller than the ICSM in the high SNR, while the standard deviations are considerable. Overall, the estimated performance of the BSCSM is better than the one of CSM; the estimated performance of the CSM is better than the one of ICSM.

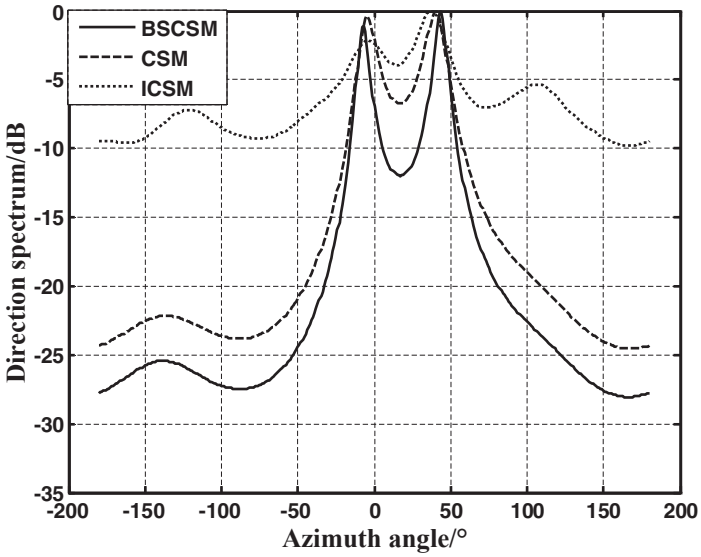


Fig. 58.5 Evaluation of two signals' azimuth angle by BSCSM, ICSM, and CSM

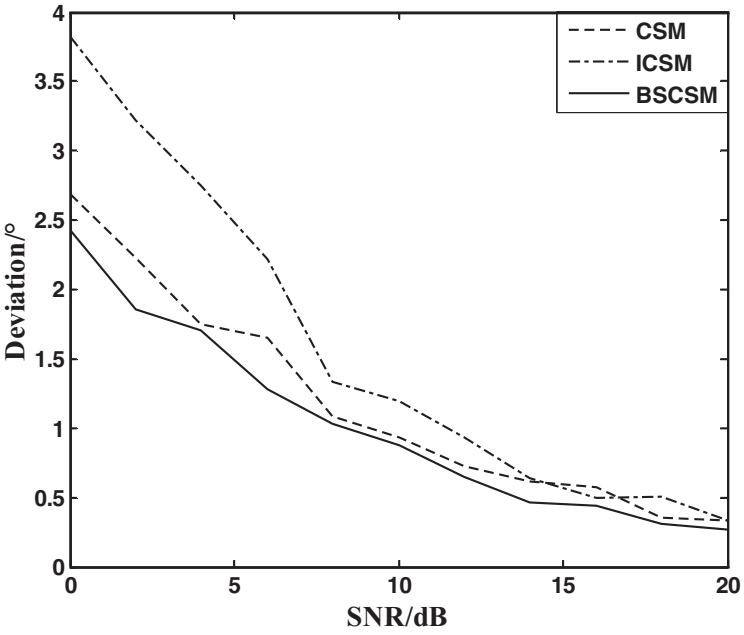


Fig. 58.6 Deviation comparison between the BSCSM, ICSM, and CSM

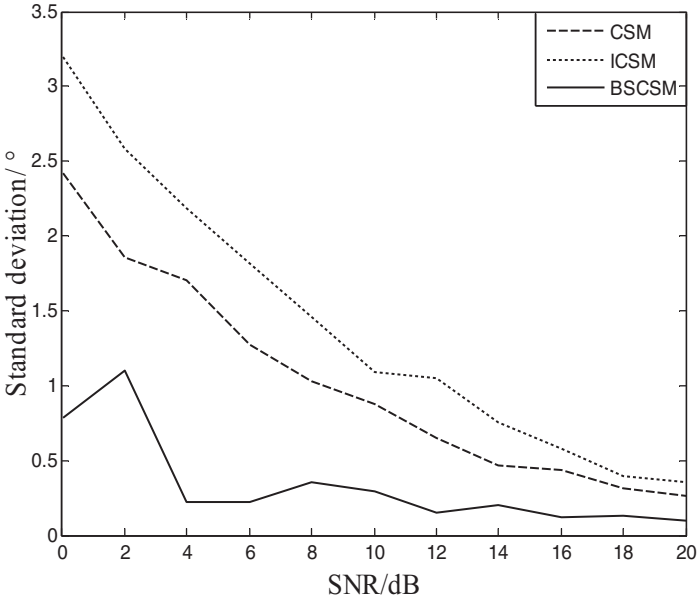


Fig. 58.7 Standard deviation comparison between the BSCSM, ICSM, and CSM

58.5 Conclusion

This chapter focuses on the DOA estimation and resolution enhancement of the two broadband sources based on the BSCSM, ICSM, and CSM with an emphasis on the four-element planar cross array. Estimated results display that the accuracy of a four-element planar cross array is not so good compared with other reports which use more array antenna elements and complex array element configurations. The four-element planar cross array has the smaller amount of calculation of space decomposition, so the estimation time is shorter. A rough estimate of the target will be enough in a battlefield where it is difficult to afford huge computation while using large array elements with complex pattern, which will lead to failure in early detection of the target. The BSCSM uses a beam-domain processing that can enhance the capacity of the array elements to provide less deviation from the estimated bias and standard deviation of two kinds of methods discussed; we can see that the BSCSM is superior to the CSM, and the CSM is superior to the ICSM. The BSCSM, ICSM, and CSM are also adopted for the DOA estimation of narrowband sources. Due to the orientation of broadband spectrum estimation by multiple narrowband average, the DOA estimation results adopting the algorithms in wideband sources are more stable than those adopting the algorithms in narrowband sources.

References

1. Shahi SN, Emadi M, Sadeghi K. High resolution DOA estimation in fully coherent environments. Prog Electromagn Res C. 2008;5(7):135–48.
2. Jia YD, Leng SL. A simple system positioning helicopter system based on passive acoustic detection. J Beijing Inst Technol. 2000;20(4):338–42.
3. Van Trees HL. Optimum array processing. New York: Wiley; 2002. S. 1435–72.
4. Naidu PS. Sensor array signal processing. New York: CRC Press LLC; 2001. 35–472.
5. Yoon YS, Kaplan LM, McClellan JH. Direction-of-arrival estimation of wideband sources using arbitrary shaped multidimensional arrays. //Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on; IEEE; 2004. ii—221–4.
6. Yang YX. Studies on beamforming and beamspace high resolution bearing estimation techniques in sonar systems. Ph. D. thesis in Northwestern Polytechnical University, Xi'an, China, June 2002. (In Chinese).
7. Yan S, Hou C. Broadband DOA estimation using optimal array pattern synthesis technique. IEEE Antennas Wirel Propag Lett. 2006;5(1):88–90. doi:10.1109/LAWP.2006.872420.
8. Wang Z, Li H, Shi KL, Zhao JW. On improving time delay estimation precision of small size array. J Northwest Polytech Univ. 2000;18(1):142–6. (In Chinese).
9. Yang YX. Beamforming and beamspace high resolution bearing estimation techniques in sonar systems. Xi'an: Northwestern Polytechnical University: Ph. D thesis, June 2002. 94–100. (In Chinese).
10. Dassot G, blanpain R, Jauffret C. Maximum likelihood estimator for magneto- acoustic localization. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing; [S.l.]: IEEE Press; 1997. 1: 495–8.

Chapter 59
Fast Segmentation Algorithm Based on Texture

Yali Qi, Guoshan Zhang and Yeli Li

Abstract The state-of-the-art image retrieval method pays more attention to the local region feature. This chapter proposes an algorithm to extract texture feature based on Laws' filter. This algorithm performs texture transformation to an image and selects the region of interest. It convolves the input image with Laws' filter mask. As to the filter mask, this chapter selects a five dimension vector and produces a 5×5 matrix as the filter kernel. The experiment shows that the algorithm has high speed and high accuracy in the segmentation process for image retrieval and confirms that it is effective for image retrieval based on the region of interest.

Keywords Region of interest · Laws' filter · Texture feature · Image retrieval

59.1 Introduction

The content-based image retrieval (CBIR) is opposed to traditional concept-based approaches. Content-based implies that the search discusses the contents of the image rather than the conceptions such as descriptions, keywords, or tags associated with the image. The term "content" refers to colors, shapes, textures, or any other features that can be extracted from the image directly. As the searches depend on the annotation quality and completeness, CBIR is desirable for image retrieval. How to search desired images in large databases quickly and accurately was a hot research

Y. Qi (✉)
The Computer Science and Technology Department,
Beijing Institute of Graphic Communication, 102600 Beijing, China
e-mail: qyl@bigc.edu.cn

School of Electrical Engineering and Automation, Tianjin University, 300072 Tianjin, China

Y. Li
Beijing Institute of Graphic Communication, 102600 Beijing, China

G. Zhang
School of Electrical Engineering and Automation, Tianjin University, 300072 Tianjin, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_59

topic in the last decade. As one of the promising approaches, the salient point and the region of interest-based image retrieval has attracted many researchers.

The salient point detection in images is very useful for image processing applications such as image compression, retrieval, object detection, or recognition. In addition, the CBIR salient point is also frequently used to represent the local properties of the image. There is a hierarchical algorithm to select the most salient points. It gives a satisfying representation of an image and makes the image retrieval systems more efficient [1]. The other method integrates the star shape of the image object using a salient point detection algorithm, which uses the Harris salient point detection and an adaptive threshold segmentation algorithm [2]. There is also a content-based image retrieval model which uses the local visual attention feature. It uses the salient image points extracted by a fast and performing speeded-up robust features (SURF) detector [3].

To find the region of interest, there is a method based on the template. It describes a way to learn the similar images given by the user, which are obtained based on the distances and spatial relationships in the similar region pairs [4]. It also remains a challenge because the low-feature similarity cannot express high-semantic similarity. Thus, the retrieval results are often far from satisfaction. To solve this problem, this chapter proposes a region-based object retrieval by using the adaptive image segmentation and generalized Hough transform [5]. In addition, A. Talib et al. proposed a semantic feature extracted from dominant colors [6]. It may reduce the effect of the image background on the image matching decision. There are some other methods to find the region of interest, such as by local difference operators [7], dominant color and local binary patterns [8], local maximum edge binary patterns [9], local Tetra pattern, local oppugnant color texture patterns [10], DLEP, block-based local binary patterns, center-symmetric local binary patterns, local edge patterns for segmentation, and local edge patterns for image retrieval methods [11]. These methods try to select the region of interest and improve the efficiency of image retrieval.

However, related work seldom considers time consumption; and some salient points or regions of interest may always not represent the most interesting regions for image retrieval. This chapter proposes an algorithm to segment the region of interest based on the Laws' texture filter.

59.2 Texture Feature

In the content-based image retrieval, the content can be expressed by color, texture, shape, and other features. The texture feature designs to quantify the perceived texture of an image. It gives us information about the spatial arrangement of intensities or color in an image. Image textures can be used to help in segmentation (image processing) or classification of images.

There are two approaches to extract the texture features, specifically, the structured approach and the statistical approach [12]. The structured approach recognizes the texture as a series of primitive texels with some regular or repeated pat-

terns. It does well in terms of analyzing artificial textures. The statistical approach recognizes the texture as a quantitative gauge of the arrangement of intensities in a region. Generally, the statistical approach is more widely used because it is easier to compute. The statistical approach to extract image texture includes three classes: the edge detection, the cooccurrence matrices, and the Laws' texture energy measure.

59.2.1 Edge Detection

The edge detection aims at identifying special points in an image whose brightness changes sharply or formally. A set of curved line segments termed edges can be gotten by those points. With the edge detection, the number of edge pixels in a specified region can be determined, the same as the characteristic of the texture complexity. Once the edges are found, we can extract the direction of the edges as a characteristic of the texture. The frequently used edge detection operator includes Robert, Prewwit, Laplacian, Marr-hidreth, Sobel, Canny, etc.

59.2.2 Co-occurrence Matrices

With the spatial relation of a similar gray tone, the numerical texture features can be extracted based on the co-occurrence matrix. These texture features can be used to represent, compare, and classify images. R. Haralick proposed 14 texture features derived from a normalized co-occurrence matrix, such as angle second moment, contrast, correlation, entropy, sum of squares, sum of average, inverse difference moment, difference of variance, etc. [13]. There are various applications of the co-occurrence matrix in the image processing; but there is a negative aspect of the co-occurrence matrix because the extracted features do not necessarily correspond to visual perception.

59.2.3 Laws' Texture Energy Measures

Kenneth Ivan Laws proposes the texture energy measures which have been used for many different applications. Firstly, the measures are computed by applying small convolution kernels to an image; secondly, the measures perform a nonlinear windowing operation. The two-dimensional convolution kernels are typically used for texture discrimination. Generally, they are generated from the following set of one-dimensional convolution kernels of length five:

Level = [+1 +4 6 +4 +1];
Edge = [−1 −2 0 +2 +1];
Spot = [−1 0 2 0 −1];
Wave = [−1 +2 0 −2 +1];
Ripple = [+1 −4 6 −4 +1].

It can generate 25 different two-dimensional convolution masks by convolving a one-dimensional kernel with another one-dimensional kernel. In the application, the convolution masks of 5×5 are used to compute the energy of texture, which can be represented by a nine element vector for each pixel.

To segment the region of interest for image retrieval, it needs high speed and high accuracy. We propose a novel segmentation algorithm for the region of interest based on Laws' texture energy measures.

59.3 Experiment

59.3.1 Segmentation Algorithm Based on Laws' Texture Filter

Based on the Laws' texture energy measure, apply a texture transformation to an image, which can be done by convolving the input image with a special filter mask. The segmentation algorithm is shown as below:

S1: As to the input image, it applies a texture transformation by convolving with a special filter mask. The filters select 5×5 matrices which obtain from the above five vectors. In our experiment, we select Edge5 × Level5.

S2: After filtering an image using Laws' texture filter, it carries out a variation of the median filtering on from the vertical and the horizontal direction, which transforms the image more smooth and separates median filtering with rectangular masks. After this step, we get two-channel images.

S3: Calculate the histogram of two-channel gray value images. Design the global threshold and select pixels from the image whose gray values fulfill the conditions.

S4: Segment the image by means of two-dimensional pixel classification, and it shows the texture region. This texture region is our region of interest.

59.3.2 Experiment Results

In this section, we evaluate the efficiency of our segment algorithm with Laws' texture filter. As the co-occurrence texture feature has seldom visual perception, we select two traditional edge detection operators to compare with our segmentation algorithm based on Laws' texture filter. The two traditional edge detection operators are Soble and Robinson operators. All experiments described in this section have been performed on a Lenovo workstation with an Intel Core 2.5Ghz CPU and a 4 GB RAM running Win8 and halcon11. The images are selected from the http://image.baidu.com/.

As shown in Fig. 59.1, the algorithm is adaptive to the image with obvious texture. If the foreground or the background is relatively simple and easily separated,

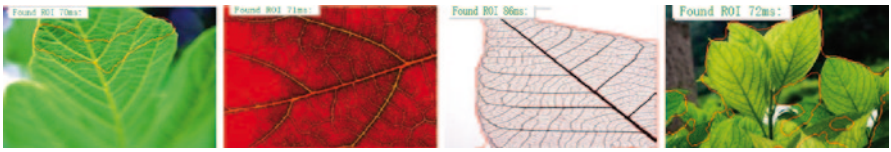


Fig. 59.1 Image with obvious texture

Fig. 59.2 Image with simple foreground and background

Fig. 59.3 Image with complex foreground and background

the algorithm can get good performance, as illustrated in Fig. 59.2. But as to the complex foreground or complex background image, the algorithm shows poor performance, as shown in Fig. 59.3.

In Fig. 59.4, the left column shows the results of the segmentation region of interest based on Laws' texture filter, the middle column shows segmentation results based on the Sobel operator, the right column shows the segmentation results based on Robinson's operator. We can see the efficiency of our algorithm by comparing it with the traditional algorithm. From the point of visual perception, segmentation algorithm proposed in this chapter can get a distinct and complete region of interest. For the traditional method, the region can be obtained, but the edge is blur and incomplete. Crucially, our algorithm features a higher speed than other algorithms obviously.

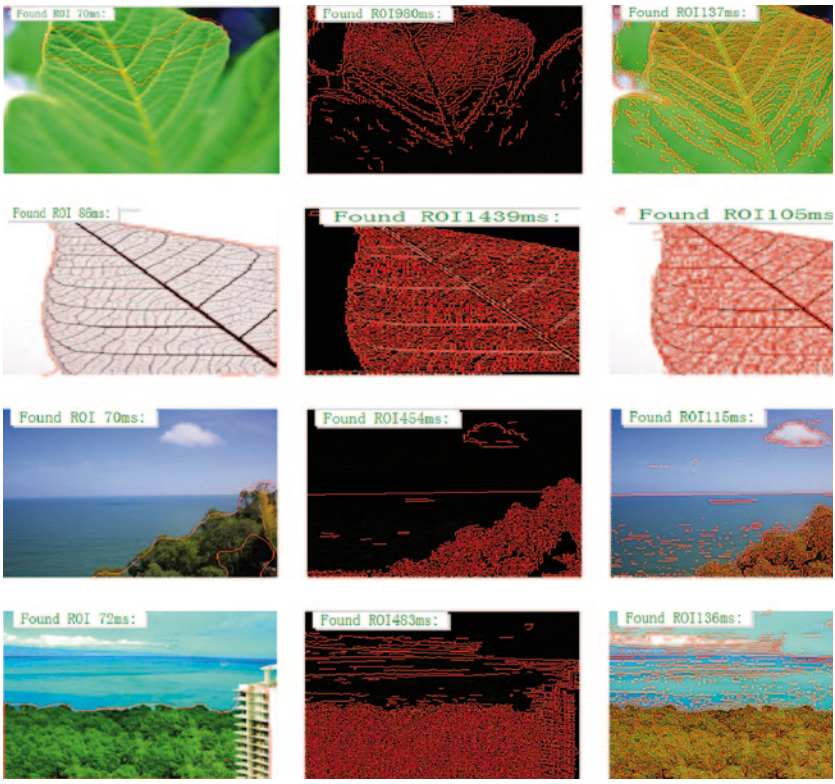


Fig 59.4 Three methods of the segmentation ROI based on texture. (*Left column:* segmentation ROI based on Laws' texture filter, *middle column:* segmentation ROI based on the Sobel operator, *right column:* segmentation ROI based on the Robinson operator)

59.4 Conclusion

In order to find the region of interest in the image retrieval, this chapter proposes an algorithm to the segmentation region of interest based on the Laws' texture filter. The experiment reveals that it shows better performance for an image with an obvious texture and a simple foreground and background. The next research is to find a new method or to improve this algorithm in order to apply it to images with a complex foreground or background. The experiments also confirm that the algorithm has higher speed and accuracy than traditional region of interest segmentation algorithms based on texture.

Acknowledgments This research is supported by the Beijing University Youth Excellence Program (YETP1467). It is supported by the BIGC Key Project (E-A-2014–11) and Program funded by the Education Commission of Beijing municipal colleges to enhance the innovation ability (TJSHG201310015016).

References

1. Tsai YH. Hierarchical salient point selection for image retrieval. Pattern Recognit Lett. 2012;33(12):1587–93.
2. Yang HY, Li YW, Li WY, Wang XY, Yang FY. Content-based image retrieval using local visual attention feature. J Vis Commun Image Represent. 2014;25(6):1308–23.
3. Liao XL, Xu HB, Zhou YC, Li KQ, Tao WB, Guo QJ, Liu LM. Automatic image segmentation using salient key point extraction and star shape prior. Signal Process. 2014;105:122–36.
4. Shen XJ, Ju SG, Cho SY, Li F. Mining user hidden semantics from image content for image retrieval. J Vis Commun Image Represent. 2008;19(3):145–64.
5. Chung CH, Cheng SC, Chang CC. Adaptive image segmentation for region-based object retrieval using generalized Hough transform. Pattern Recognit. 2010;43(10):3219–32.
6. Talib A, Mahmuddin M, Husni H, George LE. A weighted dominant color descriptor for content-based image retrieval. J Vis Commun Image Represent. 2013;24(3):345–60.
7. Subrahmanyam M, Jonathan QM. Expert content-based image retrieval system using robust local patterns. J Vis Commun Image Represent. 2014;25(6):1324–34.
8. Shrivastava N, Tyagi V. Content based image retrieval based on relative locations of multiple regions of interest using selective regions matching. Inf Sci. 2014;259(20):212–24.
9. Subrahmanyam M, Maheshwari RP, Balasubramanian R. Local maximum edge binary patterns: a new descriptor for image retrieval and object tracking. Signal Process. 2012;92(6):1467–79.
10. Jacob IJ, Srinivasagan KG, Jayapriya K. Local oppugnant color texture pattern for image retrieval system. Pattern Recognit Lett. 2014;42(1):72–8.
11. Reddy PVB, Reddy ARM. Content based image indexing and retrieval using directional local extrema and magnitude patterns. AEU—Int J Electr Commun. 2014;68(7):637–43.
12. Laws K. Textured image segmentation. Ph. D. Dissertation. University of Southern California;1980.
13. Haralick R. Statistical and structural approaches to texture. Proc IEEE, IEEE. 1979;67(5):786–804.

Chapter 60
Optimal Test Selection of Complex Electronic Systems Based on Improved Discrete Particle Swarm Optimization Algorithm

Ling Ma, Haijun Li and Xiaofeng Lv

Abstract Optimal test selection is the important content of complex electronic system testability design. This chapter establishes the mathematical model of optimal test selection and then proposes an improved discrete particle swarm optimization algorithm to provide a solution. The algorithm designs a new fitness function according to the characteristics of test selection. In order to avoid the local optimum, an inertia weight adaptive adjustment strategy based on the group's premature degree is proposed. The simulation results show that the algorithm proposed can achieve a global optimal solution fast and effectively. Optimization results meet all system requirements and can provide an effective guidance for optimal test selection of complex electronic systems.

Keywords DPSO algorithm · Testability design · Test selection

60.1 Introduction

With the performance and complexity improvement for weapon systems, fault detection and diagnosis become more and more difficult; thus, carrying out testability design is a hot topic [1]. Test optimization selection is the important content of system testability design. The purpose of optimal test selection is to solve the following two problems: (1) to identify the needed test set in the guarantee of system fault coverage and identification; (2) to select test subset meeting the system testability requirement from the alternative test set, and this subset should correspond to the lowest test cost [2].

L. Ma (✉) · H. Li · X. Lv
Naval Aeronautical and Astronautical University, 264001 Yantai, China
e-mail: miller5356@vip.sina.com

H. Li
e-mail: 19804473@qq.com

X. Lv
e-mail: lxf_00112@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control*, Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_60

Optimal test selection belongs to typical set covering and combinatorial optimization problem and is difficult to be solved. For such a problem, many papers propose the corresponding solutions [3–7]; there are methods based on Boolean logic, genetic algorithm (GA), information entropy, particle swarm optimization (PSO), and hybrid binary particle swarm optimization and genetic algorithm (HBPSOGA). But due to the inherent difficulty, neither computational efficiency nor accuracy is satisfactory. Especially with the increasing complexity of weapon equipment and scale of collection, a new effective algorithm is needed to obtain the optimal solution.

In order to improve the convergence speed and solution precision of DPSO, this chapter designs a new method of calculating particle individual fitness with consideration of the characteristics of test selection; and according to the feature that it is easy to fall into a local optimum, a strategy of inertia weight adjustment based on the degree of premature convergence is proposed; then, the improved particle swarm optimization (IDPSO) algorithm for optimal test selection is developed.

60.2 Problem Description and Modeling

Describing the binary relation of fault and test, the system fault–test dependency matrix is the basis for test selection and can be obtained by testability modeling [8]. Provided $F=\{f_1,\cdots,f_m\}$ is a set of faults in the system, $\lambda=[\lambda_1,\lambda_2,\cdots,\lambda_m]$ is its fault prior probability vector, $T=\{t_1,\cdots,t_n\}$ is the test set that can be chosen, and $C=[c_1,c_2,\cdots,c_n]$ is its test cost vector. The dependency matrix of the system can be written as the Boolean matrix $FT=\{ft_{ij}\}$, in which $i=1,2,\cdots,m,\ j=1,2,\cdots,n$. The fault–test dependency matrix is shown below [9].

$$FT=\begin{bmatrix} ft_{11} & ft_{12} & \cdots & ft_{1n} \\ ft_{21} & ft_{22} & \cdots & ft_{2n} \\ \vdots & \vdots & & \vdots \\ ft_{m1} & ft_{m2} & \cdots & ft_{mn} \end{bmatrix} \tag{60.1}$$

in which, if f_i can be observed by test t_j, then $ft_{ij}=1$, otherwise $ft_{ij}=0$.

As the chance for multiple fault occurrence of the system is very slight, it is assumed that at any time, there could be one fault occurring in the system at the most. Provided the test set is $x=[x_1,x_2,\cdots,x_n]$, $x_k=1$ shows test t_k is selected, $x_k=0$ shows test t_k is not selected, in which $k=1,\cdots,n$. T_s is the test subset selected from test set T.

The condition for fault f_i to be detected by T_s is that the row vector corresponding to f_i in the dependency matrix has at least one element 1, that is, $\bigcup_{t_j\in T_S}^{N_S} ft_{ij}=1$. In the formula, $\bigcup$ is the OR operation of Boolean variables, and N_s is the number of sets T_s. Assuming that the set of all fault can be detected by T_s, then

$$F_D = \left\{ f_i \middle| f_i \in F, \bigcup_{t_j \in T_S}^{N_S} ft_{ij} = 1 \right\} \tag{60.2}$$

Therefore, fault detection rate (FDR) can be expressed as [10]:

$$\gamma_{FD} = \frac{\sum\limits_{f_i \in F_D} \lambda_i}{\sum\limits_{f_i \in F} \lambda_i} \tag{60.3}$$

Provided a set T_{f_i} consists of all tests which can defect the fault f_i, and a set T_{f_j} consists of all tests which can defect the fault f_j, then the condition for fault f_i and f_j that can be isolated by T_s is that both fault f_i and f_j can be detected by T_s, and vectors T_{f_i} and T_{f_j} are different, that is $T_{f_i} \oplus T_{f_j} \neq 0, \forall f_i, f_j \in F_D, i \neq j$, in which $\oplus$ stands for the vector's XOR operation. If the corresponding element of the two vectors is the same for both, the result is 0, otherwise it is 1. Assuming that the set consisted of all faults can be isolated by T_s, then

$$F_I = \left\{ f_i \middle| f_i \in F_D, T_{f_i} \oplus T_{f_j} \neq 0, \forall f_j \in F_D, i \neq j \right\} \tag{60.4}$$

Therefore, the fault isolation rate (FIR) can be expressed as [10]:

$$\gamma_{FI} = \frac{\sum\limits_{f_i \in F_I} \lambda_i}{\sum\limits_{f_i \in F_D} \lambda_i} \tag{60.5}$$

The purpose of test optimal selection is to find a test set which meets the requirement of system FDR and FIR, while ensuring the minimum number and lowest cost of test under the maximum FDR and FIR. Upon above analysis, the optimal selection model is as follows [11]:

$$\begin{cases} \min \left\{ \sum\limits_{k=1}^{n} x_k c_k \right\} \\ s.t. \gamma_{FD} \geq \gamma_{FD}^* \\ \gamma_{FI} \geq \gamma_{FI}^* \end{cases} \tag{60.6}$$

In the formula (6), $\gamma_{FD}^*, \gamma_{FI}^*$ is FDR and FIR required by the system.

60.3 DPSO Algorithm

The earliest PSO version can be referenced in the literature [12]. Kennedy and Eberhart put forward the discrete particle swarm optimization (DPSO) in 1997 [13]. Each potential solution is assigned a randomized velocity, and the potential solution, which is called particle, is then "flown" through the problem space; each particle keeps track of its coordinates in the problem space, which are associated with the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called P_{bestid}. Another "best" value that is tracked by the global version of the particle swarm optimizer is the overall best value and its location, which is obtained so far by any particle in the population. This location is called G_{bestid}. The velocities and positions of particles in DPSO are updated as follows [12]:

$$\begin{cases} v_{id}^{t+1} = \omega v_{id}^{t} + c_1 r_1^t (P_{bestid}^t - x_{id}^t) + c_2 r_2^t (G_{bestid}^t - x_{id}^t) \\ x_{id}^{t+1} = \begin{cases} 1, \ \delta_{id}^{t+1} < sigmoid(v_{id}^{t+1}) \\ 0, \ \delta_{id}^{t+1} \geq sigmoid(v_{id}^{t+1}) \end{cases} \\ sigmoid(v_{id}^{t+1}) = 1/(1+exp(v_{id}^{t+1})) \end{cases} \tag{60.7}$$

In the above formula, x_{id}^t and v_{id}^t are the position and velocity of the particle in the tth iteration, ω is the inertia weight, c_1 and c_2 are the acceleration coefficients, r_1^t and r_2^t are the random values between (0,1) in the tth iteration, P_{bestid}^t is the best position that the particle ever had in the tth iteration, G_{bestid}^t is the best position that the group particles ever had in the tth iteration, and δ_{id}^{t+1} is a random value between (0,1).

60.4 IDPSO Aiming at Optimal Test Selection

60.4.1 Population Initialization

The particles of DPSO consist of a binary code with a length of *n,* each of which corresponds to a test set. When the algorithm starts, population initialization $T(0) = \{T_1^0, T_2^0, \cdots, T_M^0\}$, in which M is population size, $T_j^0 (j = 1, 2, \cdots, M)$ is the jth particle, and its corresponding binary code is $T_j^0 = [t_{1j}^0, t_{2j}^0, \cdots, t_{nj}^0]$, in which $t_{ij}^0 = 1$ means the ith test of test set T_j^0 is chosen, otherwise T_j^0 is not chosen; $i = 1, 2, \cdots, n$.

60.4.2 Fitness Evaluation

Evolving to generation k, current population is $T(k) = \{T_1^k, T_2^k, \cdots, T_M^k\}$. In order to evaluate the fitness of particle $T_j^k (j = 1, 2, \cdots, M)$, the following method of calculating the fitness of T_j^k is designed, and the specific steps are described as follows:

Step 1: Record T_j to the selected test set; F_j is the fault set detected by T_j; use formula (3) to calculate the $\gamma_{FD}(T_j)$ of test set T_j. If $\gamma_{FD}(T_j) \geq \gamma_{FD}^*$, go to step 5, otherwise go to step 2.

Step 2: Get test set $\overline{T_j} = T - T_j$ from T_j and get fault set $\overline{F_j} = F - F_j$ from F_j.

Step 3: Calculate the ratio of each test from test set $\overline{T_j}$ which is not selected: p/c, in which p is the sum of fault probability which can be tested by the corresponding test in $\overline{F_j}$; c is the test cost of the corresponding test. Select the test with maximum ratio to add to test set T_j, and add the faults to fault set F_j.

Step 4: Calculate FDR $\gamma_{FD}(T_j)$ of test set T_j; if $\gamma_{FD}(T_j) \geq \gamma_{FD}^*$, go to step 5, otherwise update $\overline{T_j}$ and $\overline{F_j}$, and go to step 2.

Step 5: According to the model shown in formula (6), define the fitness function shown using the method of penalty function and calculate the fitness of T_j:

$$Fitness = \frac{C}{\sum_{t_j \in T_j} x_j c_j + \sum_{t_j \in T} x_j c_j} - \alpha max(0, \gamma_{FI}^* - \gamma_{FI}) \tag{60.8}$$

in which c_j expresses the test cost of test t_j; C and α are constants. When test set T_j satisfies the FIR index, the individual fitness is decided by the first part, otherwise the individual fitness is decided by the second part, to ensure that the individual chosen by the fitness function is the best individual in the current population.

60.4.3 Adaptive Inertia Weight Adjustment

As to the traditional DPSO algorithm, inertia weight adjustment strategy in linear decreasing is widely used. The strategy can make the population get better global search ability in early optimization. With the decrease of inertia weight, it contributes to the convergence of global extreme point; however, in this strategy, all particles use the same weight value, which is not conducive to the diversity of the population; at the same time, constant decline also reduced the solving efficiency. According to the relationship of premature convergence degree of swarm and particle fitness value, a new adaptive inertia weight adjustment strategy is raised. Compared with the adaptive inertia weight adjustment strategy based on particle space [14], the proposed strategy is more advanced, which uses different adaptive operations for different particles, making the group always maintain the diversity of inertia weight.

The premature convergence degree of the particle swarm can be evaluated by Δ; the smaller Δ, the more premature the convergence of the particle swarm.

$$\Delta = \left| f_g - f_p \right| \tag{60.9}$$

in which f_g is the fitness value of the best global particle, f_{ag} is the average value of fitness among all current particles, f_p is the fitness value better than f_{ag}; this chapter takes the intermediate value of f_g and f_{ag}.

For particles with fitness value f_i, the adjustment strategy of inertia weight is as follows:

1. When $f_i > f_p$, the particle is better in the population, which is close to the global optimal value, and the inertia weight should be smaller accordingly, to strengthen the local search.

$$\omega' = \omega_{min} + (\omega_a - \omega_{min}) \bullet \left| \frac{f_i - f_p}{\Delta} \right| \tag{60.10}$$

In which ω_{min} is the minimum value of ω, and ω_a is the intermediate value of the ω value range $[\omega_{min}, \omega_{max}]$.

2. When $f_{ag} \le f_i \le f_p$, the particle is a general particle in the population, which has a good ability of local optimization and global optimization, and the inertia weight uses the traditional linear decreasing strategy for adaptive adjusting.

$$\omega' = \omega_{min} + (\omega_{max} - \omega_{min}) \bullet \frac{MaxIter - CurIter}{MaxIter} \tag{60.11}$$

In which ω_{max} is the maximum value for ω, $CurIter$ and $MaxIter$ are the current iterations and maximum iterations of algorithm.

3. When $f_i < f_{ag}$, the particle is poor in the population, which should be given greater inertia weight.

$$\omega' = \omega_{max} - \frac{1}{1 + k_1 \bullet exp(-k_2 \bullet \Delta)} \tag{60.12}$$

In which k_1 is the control parameter, k_2 is the adjusting parameter, which meet the condition of $k_1 > 1, k_2 > 0$ here.

60.5 Simulation and Verification

To verify the effectiveness of the proposed algorithm, an example is introduced to prove the performance of IDPSO. The operating environment is as follows: CPU Intel Core I3 3.30 GHz, RAM 2 GB, Matlab 8.0.0.783.

This example is illustrated from a super-heterodyne receiver [7], in which there are 22 prior faults and 36 alternative tests. The cost of each test is equal to one, and the fault prior probability is as follows: [0.00185, 0.00923, 0.185, 0.00185, 0.00185, 0.00923, 0.00185, 0.00923, 0.185, 0.185, 0.185, 0.00185, 0.00923, 0.185, 0.00923, 0.00185, 0.00923, 0.00185, 0.00185, 0.00185, 0.00185, 0.00185]. Testability index requires that FDR and FIR are no less than 95 %.

Using the IDPSO method to solve, set parameter values as follows: $Popsize$=60, ω_{max}=1.2, ω_{min}=0.4, N_{max}=200, c_1=c_2 = 1.4962, C = 10, α= 0.5 .

The optimal result is [0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0]; i.e., the test set is $\{t_8, t_{26}, t_{28}, t_{31}, t_{32}, t_{34}\}$, actual FDR is 96.86%, FIR is 96.57%, test cost is 6, while the test cost concluded by reference [7] is 7.

For comparison, when DPSO is used to solve this problem, set the parameter values as follows: $Popsize$=60, ω_{max}=1.2, ω_{min}=0.4, N_{max}=200, c_1=c_2 = 1.4962 . The simulation result of DPSO is [0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0]; i.e., the test set is $\{t_2, t_8, t_{26}, t_{29}, t_{31}, t_{32}, t_{34}\}$, actual FDR is 96.86%, FIR is 97.71%, and the test cost is 7.

Figure 60.1 shows the fitness variation tendency of IDPSO and Fig. 60.2 shows the fitness variation tendency of DPSO. Each iteration curve is completed after 20 times of calculation, with each calculation obtained after 200 iterations.

As can be seen from Fig. 60.1 and Fig. 60.2, the IDPSO method can satisfy system requirements and gain a global optimal solution around 30 times iteration convergence basically, and the success rate is quite high. Although the DPSO method can also satisfy system requirements, it only gains near-optimal solutions and the speed of convergence is slower than IDPSO.

Compared with the HBPSOGA method in reference [7], FIR has decreased, but the test cost is reduced under the premise of meeting testability requirement, thus obtaining a global optimal solution.

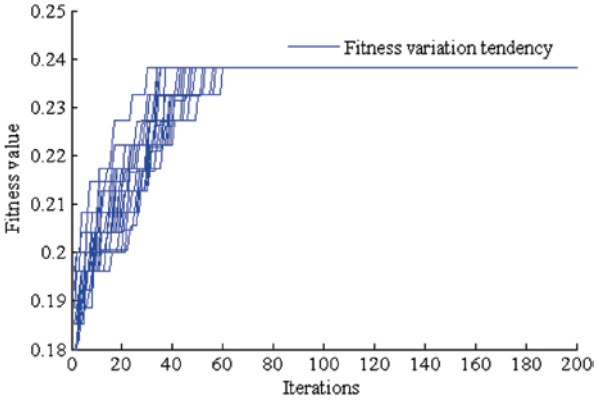


Fig. 60.1 Fitness variation tendency of IDPSO

60.6 Conclusion

With the increasing scale of the system, the correlation between faults and tests becomes more and more complex, and the test selection problem has become increasingly apparent, thus improving optimization efficiency, and accuracy becomes

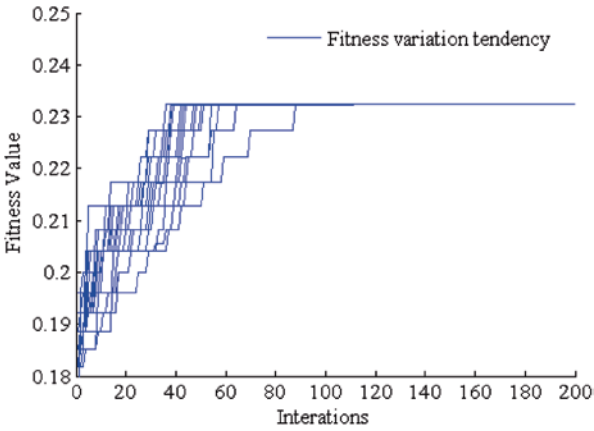


Fig. 60.2 Fitness variation tendency of DPSO

particularly important. Based on DPSO and combined with the characteristics of test selection, IDPSO designs a new method of calculating particle fitness, which provides IDPSO a better solution. And to adjust inertia weight according to the degree of particle swarm premature, not only avoid the algorithm going into premature convergence and local optimum, but also accelerate the convergence rate. The simulation result shows that, for the test selection problem, the calculation precision and convergence speed of IDPSO are better than those of HBPSOGA.

References

1. US Department of Defense. MIL-STD-2165 A. Military standard testability program for systems and equipments; 1993:1–10.
2. US Department of Defense. MIL-HDBK-2165. Testability Handbook for systems and quipment; 1995:1–16.
3. Peng Y, Jing Q, Guanjun L. The test selection algorithms based on boolean logic. J Test Measurement Technol. 2007;21(5):386–390.
4. Golonek T, Rutkowski J. Genetic-algorithm-based method for optimal analog test points selection. IEEE Trans Circuits Syst. 2007;54(2):117–121.
5. Pattipati K, Alexandridis M. Application of heuristic search and information theory to sequential fault diagnosis. IEEE Trans Syst Man Cybern. 1990;20(4):872–887.
6. Ronghua J, Houjun W, Bing L. Test selection based on binary particle swarm optimization. J Electron Meas Instrum. 2008;22(2):11–15.
7. Xixiang C, Jing Q, Guanjun L. Optimal test selection based on hybrid BPSO and GA. Chin J Scientific Instrum. 2009;30(8):1674–1679.
8. Deb S, Pattipati KR, Raghavan V, et al. Multi-signal flow graphs: a novel approach for system testability analysis and fault diagnosis. IEEE AES Mag. 1995;32(5):14–25.
9. Azam M, Tu F, Pattipati KR, et al. A dependency model-based approach for identifying and evaluating power quality problem. IEEE Trans Power Deliv. 2004;19(3):1154–1166.
10. Zhong T, Junyou S. System testability design, analysis and verification. Beijing: Beijing University of Aeronautic and Astronautics Press; 2003. p. 98–105.
11. Jing Q, Guanjun L, Peng Y, et al. Equipment testability modeling and design technology. Beijing: Science Press; 2012. p. 63–75.

12. Kennedy J, Eberhart RC. Particle swarm optimization. Proceedings of the IEEE International Conference on Neural Networks, IEEE; 1995. p. 1942al Ne.
13. Kennedy J, Eberhart RC. A discrete binary version of the particle swarm algorithm. Proceedings of the IEEE World Congress on Systemics, Cybernetics and Informatics, IEEE; 1997. p. 4104swarm.
14. Wenzhong G, Guolong C. Fuzzy self-adapted particle swarm optimization algorithm for travelling salesman problems. Comput Sci. 2006;33(6):161–163.

Chapter 61
An Improved Shuffled Frog Leaping Algorithm to Optimize the Parameters of PID

Yueting Liu

Abstract As to the nonlinear, time-varying, and large delay control system, the conventional proportional-integral-derivative (PID) control effect is not ideal; thus, PID parameters based on shuffled frog leaping algorithm (SFLA) are presented. Concerning the problems of the SFLA such as slow convergence rate and local optimality, we propose the improved shuffled frog leaping algorithm (ISFLA).This algorithm uses reverse selection mechanism in the evolutionary process so as to keep the diversity of population and improve the ability of evolution algorithm. Introducing the linear decreasing adaptive inertia weight to correct the poor frog update strategy can balance the global search and local search. The simulation results of experiments on the two classical control system show that the ISFLA, when compared with SFLA and particle swarm optimization (PSO), can balance the global search and local search with a smaller number of iterations and stronger optimization ability and is more suitable for the tuning of the PID parameters optimization.

Keywords PID parameters · SFLA · Reverse selection mechanism · Inertia weight

61.1 Introduction

The proportional-integral-derivative (PID) controller is widely used in chemical industry, electric power and other industrial production control because of advantages of simple structure, strong robustness, high reliability, and easy operation [1]. The traditional PID controller which tends to get the ideal control effect must carry on the tuning parameters first, namely the proportion coefficient (P), integral (I), and differential (D); therefore, the parameters of PID controller tuning and optimization have become an important research field.

The most actual control processes are nonlinear, time varying, and with large delay. It is difficult to establish an accurate control model and get an ideal control

Y. Liu (✉)
School of Electronics and Information Engineering, Lanzhou University of Arts and Science, 730000 Lanzhou, China
e-mail: liuyueting996@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_61

effect by the traditional PID controller [2]. In recent years, many artificial intelligence algorithms, such as genetic algorithm (GA), simulated annealing algorithm (SA), and particle swarm optimization (PSO), have been successfully applied to the PID parameter optimization [3–6]; however, GA is of low efficiency in the interrelated parameter optimization problem that it is easy to fall into local optimum; the convergence speed of SA is slower; parameters of PSO algorithm are so many that it is easy to fall into local optimum in the late algorithm. Aiming at these problems, the SFLA with the advantages of GA and PSO algorithm is hereby presented. The improved shuffled frog leaping algorithm is used to optimize PID parameters. The simulation results show ISFLA can be better applied to the project because it may improve the convergence speed and optimization of high precision.

61.2 SFLA

Please randomly generate F frogs (solution) to form the initial group. The ith frog can be expressed as $X_i = (x_{i1}, x_{i2}, \cdots, x_{id})$. For the d -dimensional problem, the frogs are in descending order according to the individual fitness. The group is divided into s subgroups with each subgroup containing n frogs. Meet the relation of $F = s \times n$. In each subgroup, the best fitness frog is recorded as X_b, the worst fitness frog recorded as X_w, and the group of the best fitness frog recorded as X_g. In each subgroup, the local search is completed. The worst frog updates strategy [7, 8].

$$D_i = rand(\quad) \cdot (X_b - X_w) \tag{61.1}$$

$$X_w = X_w(\text{The current position}) + D_i \quad \left(-D_{\max} \le D_i \le D_{\max}\right) \tag{61.2}$$

In the formula, D_i is the moving distance of the component i; $rand\ ()$ is a random number 0–1; $D_{\max}$ is the maximum distance that the frog allows to move.

After the update, if the solution X_w is better than the original solution X_w (current position), the original population of solution is replaced; otherwise, please use X_g instead of X_b; then repeat formula (61.1), (61.2). If there is no improvement, we randomly generate a new solution instead of the original X_w.

61.3 ISFLA

61.3.1 Update of the Moving Step

As to the reference update strategy of PSO algorithm [9], we improve the step of the worst frog's update in the SFLA. The inertia weight ω is introduced to balance the global and local search ability. Formula (61.3) instead of formula (61.1) will improve the step of the worst frog.

$$D_i = \omega D_i' + rand(\ )\cdot(X_b - X_w) \tag{61.3}$$

In the formula, ω is the inertia weight and D_i' is the step of the previous worst frog.

61.3.2 Reverse Selection Mechanism

The smaller the individual fitness value of the frog, the greater the probability to be selected. It can lead to the optimal solution of subgroup focus small fitness, which destroys the population diversity. The algorithm is easy to fall into a local optimum. In order to improve the evolutionary ability of the SFLA, we should maintain the diversity of population by introducing the negative selection mechanism. The smaller the reciprocal of the fitness value, the more likely it is selected. Formula (61.4) achieves to maintain population diversity.

$$P_i = 1 - \frac{\dfrac{1}{f(x_i)}}{\sum_{i=1}^{F} \dfrac{1}{f(x_i)}} \tag{61.4}$$

In the formula, P_i is the ith frog probability of the selection; $\dfrac{1}{f(x_i)}$ is the reciprocal of the fitness value of the ith frog; the formula $\sum_{i=1}^{F} \dfrac{1}{f(x_i)}$ is the reciprocal of the fitness value of the F frogs.

61.3.3 Linear Adjustment of the Inertia Weights ω

The inertia weight ω can measure the movement trend of the frog. The larger ω is conducive to jump out of the local optimal solution and avoids falling into premature. The smaller ω is conducive to complete a more accurate local search in the vicinity of the global optimal solution so as to find the global optimal solution quickly. We can use formula (61.5), the weight ω is linearly decreasing when the algorithm just enters the search. When the algorithm is close to the optimal solution, ω will gradually decrease.

$$\omega_i(t) = \omega_{\max} - (\omega_{\max} - \omega_{\min})\cdot\frac{t}{it\cdot tit} \tag{61.5}$$

In the formula, $\omega_i(t)$ is the ith frog inertia weight of the tth ; $\omega_{\max}$ and $\omega_{\min}$ are, respectively, the maximum and minimum inertia weights; $t = it_j \times tit_k, it_j$ is the iteration of the current subgroup; it is the total iteration of the subgroup; tit_k is the

iteration of the current mix group; tit is the total iteration of the mix group; j is the integer between 1 and it; and k is the integer between 1 and tit.

61.4 Optimization of PID Parameters by ISFLA

61.4.1 PID Controller

As to the unit's negative feedback control system, continuous PID controller can be expressed as follows:

$$G(s) = \frac{u(s)}{e(s)} = k_p + \frac{k_i}{s} + k_d s \tag{61.6}$$

The PID controller is a proportional, integral, and differential calculation for the system error, which constitutes the system control. In each sampling cycle of the control process, k_p, k_i, and k_d of each group will be seen as a frog of the ISFLA. With the help of the ISFLA optimizing ability, the frogs make the interaction to look for the best. With a group of the best k_p, k_i, and k_d, we can achieve the desired output.

61.4.2 Fitness Function

To achieve the optimum control quality, we must have a good set of optimization parameters k_p, k_i, and k_d. We define formula (61.7) as the fitness function for optimization and evaluation.

$$F_i(K) = \int_0^\infty (m_1 \mid e(t) \mid + m_2 u^2(t))dt + m_3 t_r \tag{61.7}$$

In the formula, $F_i(K)$ is the fitness value of the ith frog; t_r is the rise time; m_1, m_2, and m_3 are the weights; and K is $[k_p, k_i, k_d]$ of the frog. t_r and $e(t)$ are the factors of the fitness function.

61.4.3 Process of ISFLA Optimizing PID Parameter

The thought of ISFLA optimizing PID parameter is the optimal solution in all values of the unknown parameters k_p, k_i, and k_d portfolio. That is the smallest fitness function of formula (61.7). The flow chart for the ISFLA-PID is shown in Fig. 61.1.

Step 1: parameter initialization. Determine the range of the parameters k_p, k_i, and k_d; the number of frog in population is F; the number of subgroups is s;

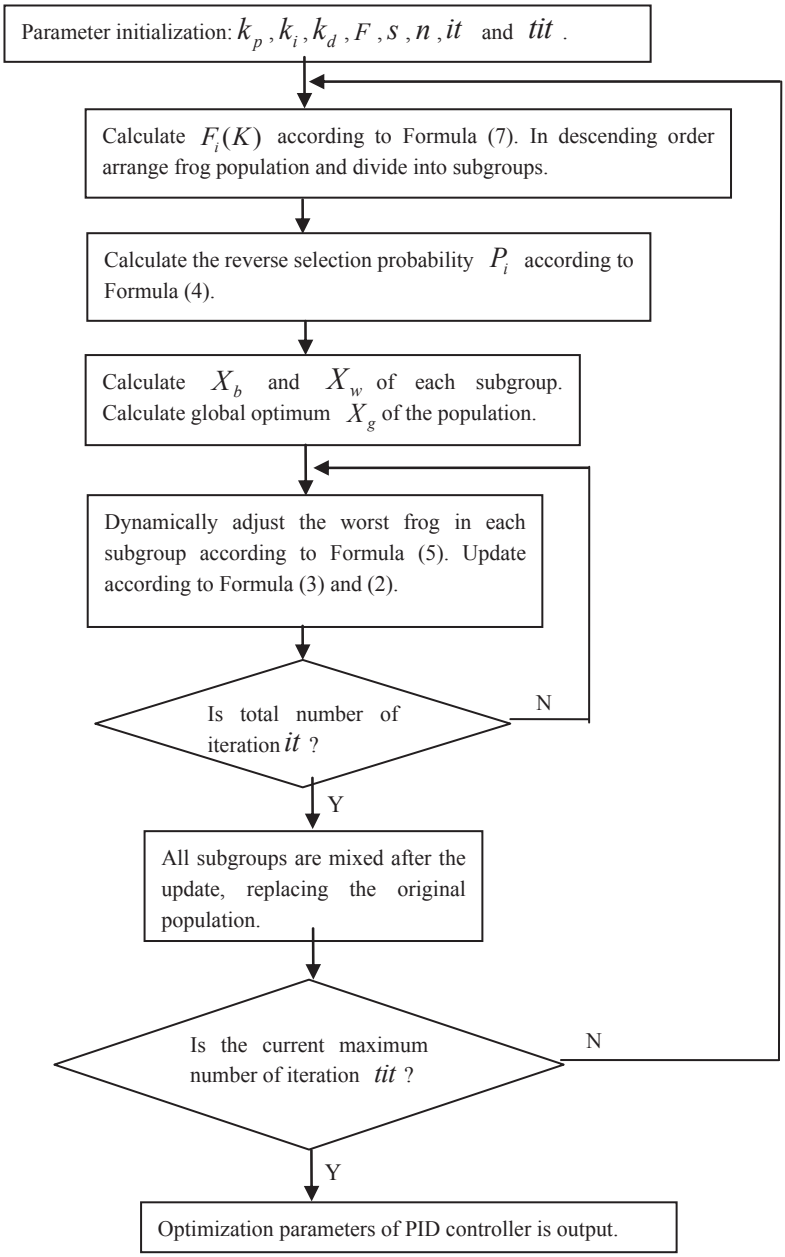


Fig. 61.1 Flow chart for the ISFLA-PID

the number of frogs in each subgroup is n; it is the total iteration of the subgroup; tit is the total iteration of the mix group.

Step 2: Calculate the fitness value $F_i(K)$ of each frog according to formula (61.7). The frog population is arranged in descending order according to $F_i(K)$ and divided into subgroups.

Step 3: Calculate the reverse selection probability P_i of each frog according to formula (61.4) so as to start the subgroup search from the worse individual fitness value.

Step 4: Calculate X_b and X_w of each subgroup. Calculate the global optimum X_g of the population. Dynamically adjust the worst frog in each subgroup according to formula (61.5). Update the operation according to formulas (61.3) and (61.2) until setting the total number of iterations. All subgroups are mixed upon updating; replace the original population.

Step 5: If the current maximum number of iterations reaches a predetermined setting, the iteration will stop. The information of optimal fitness value is put out. The algorithm ends, and then go to step 6; otherwise, go to step 2.

Step 6: Optimization parameters of the PID controller are put out.

61.5 Simulation and Experiment

To verify the performance of PID parameters, we select two control system models. We compare them with PSO-PID [5] and SFLA-PID.

The transfer functions of three systems are:

$$(1)\ G_1(s) = \frac{1}{(1+s)(1+s/8)^3}$$

$$(2)\ G_2(s) = \frac{0.33}{(1+18.5s+134s^2)} e^{-18.25s}$$

We use the parameters [10] as proposed in the algorithm performance test. The total number of Frog $F = 200$. The subgroup number $s = 20$. The total iteration of the subgroup $it = 10$. The total iteration of the mix group $tit = 500$. The range of the parameters $k_p, k_i,$ and k_d are $[0,6],[0,4],$ and $[0,4]$.

61.5.1 High-Order System

To optimize high-dimensional systems, the transfer function $G_1(s)$ is selected as a test object. The curve in Fig. 61.2 is the iterative process of the fitness function. The curve in Fig. 61.3 is the closed-loop step response and corresponds to $G_1(s)$. PID controller parameters and the PID performance simulation are shown in Table 61.1.

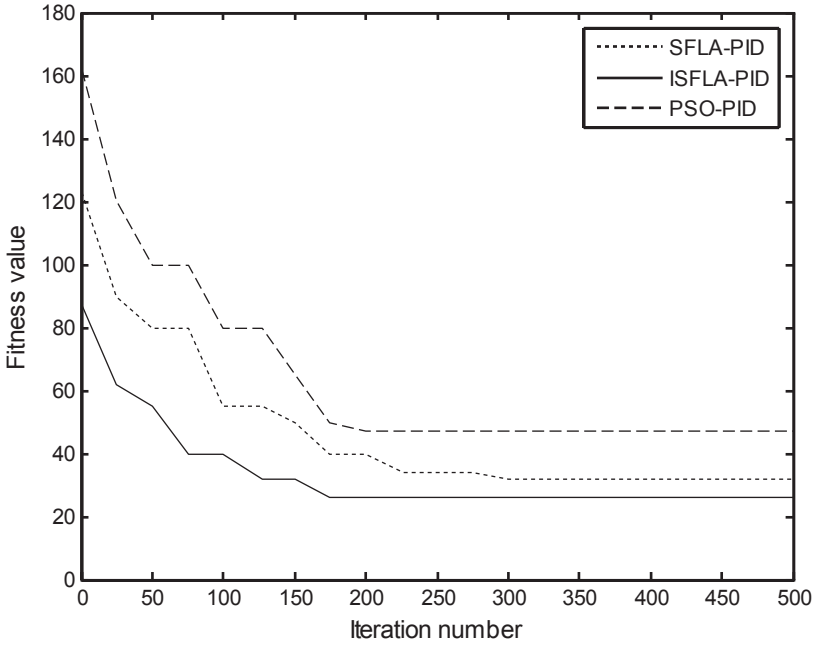


Fig. 61.2 Iterative process of the fitness function $\left(G_1(s) \right)$

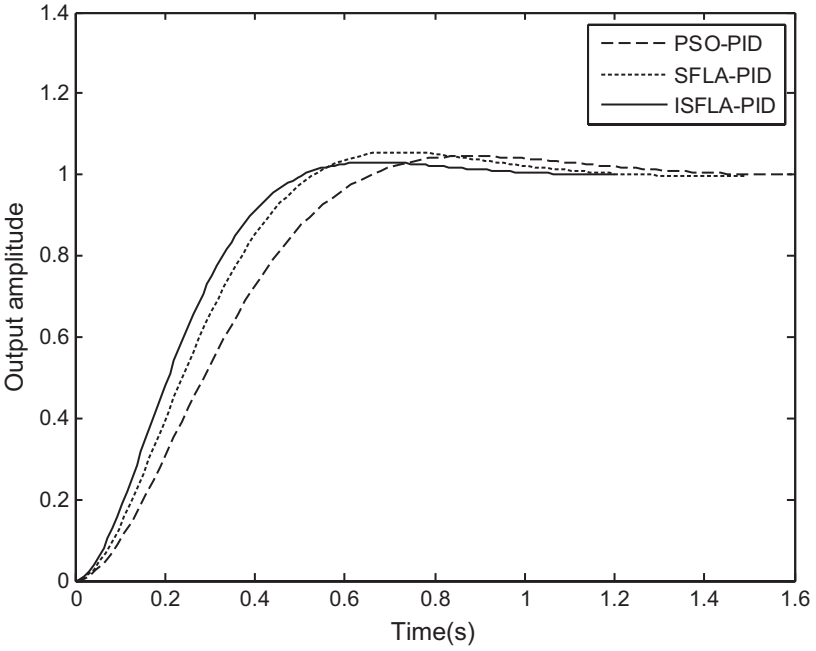


Fig. 61.3 The closed-loop step response of $G_1(s)$

Table 61.1 PID controller parameters and performance indicators $(G_1(s))$

	PID parameter tuning			PID performance	The factors of fitness function	
	k_p	k_i	k_d	Overshoot (%)	Rise time (s)	System error
PID	5.873	3.0837	3.875	6.87	5.04	0.58
PSO-PID	5.684	2.805	2.813	2.42	0.84	0.46
SFLA-PID	5.416	2.706	2.689	2.21	0.71	0.37
ISFLA-PID	4.943	2.461	2.486	1.51	0.46	0.19

61.5.2 Delay System

To optimize delay systems, the transfer function $G_2(s)$ is selected to run 20 times. The curve in Fig. 61.4 is the iterative process of the fitness function. The curve in Fig. 61.5 is the closed-loop step response and corresponds to $G_2(s)$. PID controller parameters and the PID performance simulation are shown in Table 61.2.

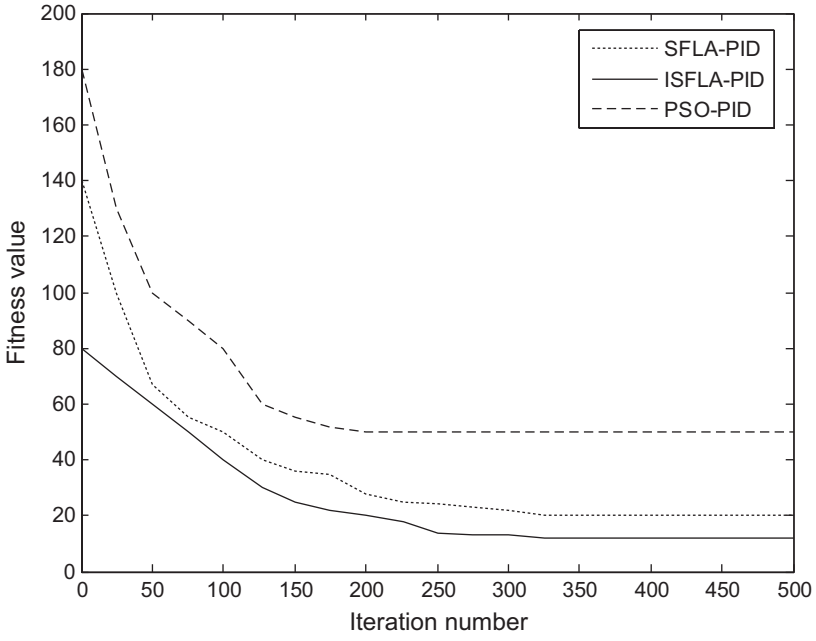


Fig. 61.4 Iterative process of the fitness function $(G_2(s))$

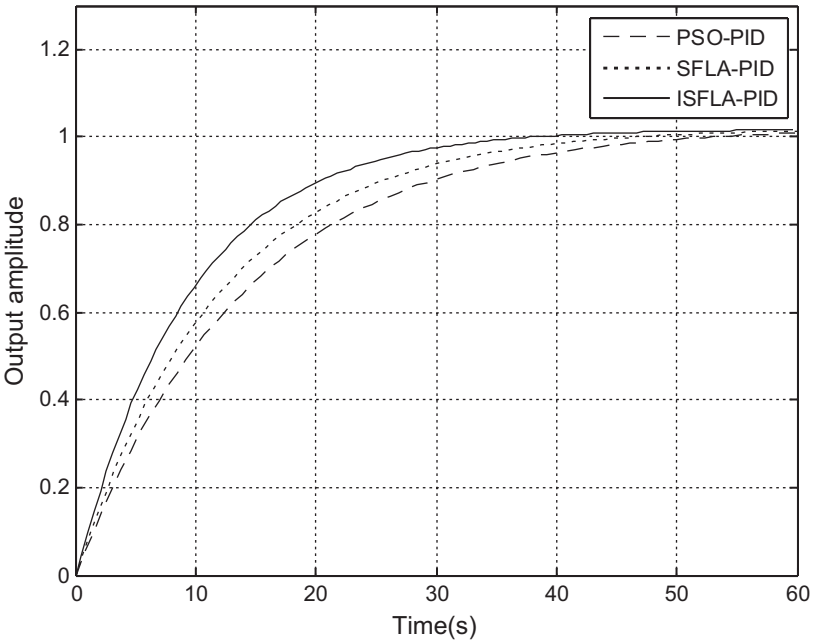


Fig. 61.5 The closed-loop step response of $G_2(s)$

Table 61.2 PID controller parameters and performance indicators $(G_2(s))$

	PID parameter tuning			PID performance	The factors of fitness function	
	k_p	k_i	k_d	Overshoot (%)	Rise time (s)	System error
PID	5.974	3.948	3.986	7.97	5.87	0.75
PSO-PID	5.684	2.805	2.813	2.42	0.84	0.46
SFLA-PID	5.416	2.706	2.689	2.21	0.71	0.37
ISFLA-PID	4.943	2.461	2.486	1.51	0.46	0.19

61.6 Conclusion

For the nonlinear, time-varying, large delay control system, the conventional PID control effect is not ideal. PID parameters optimization based on the ISFLA is presented. The algorithm makes improvement from the following three aspects: (1) update the correction strategy of the SFLA using the PSO algorithm to improve the efficiency of local search; (2) use the reverse selection mechanism in the evolutionary process so as to keep the diversity of population and improve the ability of evolution algorithm; (3) according to the actual location, select inertia weight corresponding to balance the global search and local search. The simulation results of

experiments show that the ISFLA, compared with the SFLA and PSO, can balance the global search and local search with a smaller number of iterations and stronger optimization ability.

Acknowledgments This paper is supported by the general project plan to promote the scientific research ability by Lanzhou University of Arts and Science (2012YBTS04).

References

1. Wang W, Zhang JT, Chai TY. A survey of advanced PID parameter tuning methods. Acta Automatica Sinica. 2000;26(3):347–55. (In Chinese).
2. Li W. Survey of intelligent control and its application. J Chongqing Univ Posts Telecommun (Natural Science). 2006;26(3):347–55. (In Chinese).
3. Kwok DP, Sheng F. Genetic algorithm and simulated annealing for optimal robot arm PID control. Proceedings of the IEEE Conference on Evolution of Computer. Orlando. (Hong Kong Polytech); 1994. p. 707–13.
4. Wang Q, Ma L, Zhang WY. Design for 2-DOF PID controller based on hybrid genetic algorithm and its application. Control Decis. 2001;16(2):195–202. (In Chinese).
5. Wang JS, Wang JC, Wang W. Self tuning of PID parameters based on particle swarm optimization. Control Decis. 2005;20(1):73–6, 81. (In Chinese).
6. Wang QL, Gao XZ, Wang CH. Swarm immune algorithm for PID controller self-turning. Syst Eng-Theory Pract. 2010;30(6):1062–66. (In Chinese).
7. Amiri B, Fathian M, Maroosi A. Application of shuffled frog-leaping algorithm on clustering. Adv Manuf Technol. 2009;45(1):199–209.
8. Han Y, Cai JH, Zhou GG. Advances in shuffled frog leaping algorithm. Comput Sci. 2010;37(7):16–9. (In Chinese).
9. Wang JW, Wang DW. Experiments and analysis on inertia weight in particle swarm optimization. J Syst Eng. 2005;20(2):194–8. (In Chinese).
10. Elbeltagi E, Hegazy T, Grierson D. Comparison among five evolutionary-based optimization algorithms. Adv Eng Inform. 2005;19(1):43–53.

Chapter 62
A Simple and Fast Edge Finding Algorithm for Generating an RWG Function

Fengtao Wu, Chonghua Fang and Qifeng Liu

Abstract The search pointer array data structure is presented according to the index characteristic of the edge node on the basis function; thus, a new edge finding algorithm is developed with reduced numerical complexity, of the order $O(N)$, instead of the complexity $O(N^2)$, for classical direct search, or $O(N\log(N))$ for the Adelson-Velsky and Landis (AVL) tree. Numerical experiments show that the new algorithm features fast speed, simplicity, and less memory requirement, which meets the demand of large-scale computational electromagnetic problems as solved by a multilevel fast multipole method (MLFMM).

Keywords AVL tree · Basis function · Computational electromagnetic · MLFMM

62.1 Introduction

When Rao–Wilton–Glisson (RWG) basis functions may be applied to discretize the surface integral equation in electromagnetic problems, the interior edges are needed [1]; however, the mature automatic mesh generation software, such as Patran, Ansys, FEMAP, etc., only provided the finite element connectivity file, an example of which is illustrated in Fig. 62.1. This file comprises a node list containing the Cartesian positions of all triangle nodes and a face list that assembles individual triangles by referencing three vertexes from the node list. The example in Fig. 62.1 contains six nodes and four faces. It is necessary to convert this geometric description into an enumeration of RWG basis functions. The enumeration process must be performed efficiently, or else it can become a bottleneck in simulations especially for large-scale computational electromagnetic (CEM) problems solved by a multilevel fast multipole method (MLFMM) with a large number of unknowns [2–3].

An MLFMM with computational complexity of $O(N\log(N))$ and memory requirement of $O(N\log(N))$ is introduced in CEM to accelerate the matrix vector multiplication, which has greatly increased the capacity of method of moment

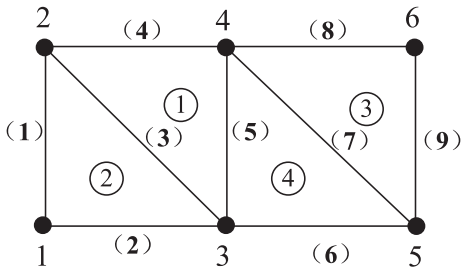
F. Wu (✉) · C. Fang · Q. Liu
Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Centre, 430000 Wuhan, China
e-mail: wk9183@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_62

Fig. 62.1 Example connectivity list in a mesh file

```
Coordinates
    1    0.0    0.0    0.0
    2    0.0    1.0    0.0
    3    1.0    0.0    0.0
    4    1.0    1.0    0.0
    5    2.0    0.0    0.0
    6    2.0    1.0    0.0
Facet
    1    4    2    3
    2    1    3    2
    3    5    6    4
    4    4    3    5
```

Fig. 62.2 Simple flat plate geometry with node numbering, facet numbering, and edge numbering



(MOM) to compute large-scale electromagnetic problems and been used in a variety of applications. Where N is the number of unknowns, the parallel MLFMM can solve electromagnetic problems with unknowns up to 85,148,1600, and fast algorithm should be applied to find the interior edge for such large CEM problems (Fig. 62.2) [4].

The simplest and easiest algorithm for interior edge finding is shown as below:

Algorithm 1:

```
1  Foreach triangular
2    Foreach edge of triangular
3      Foreach gotten interior edge
4        Check if the edge of triangular already existed in the interior edge list
5        If yes, break
6        If no, add to the interior edge list
7      End
8    End
9  End
```

Table 62.1 Time for computing interior edges and MLFMM (perfect metal sphere)

Unknowns	Algorithm 1(s)	Algorithm 2(s)	MLFMM(s)
11664	0.65625	0.0000000	55.34375
46656	9.31250	0.0156250	242.1094
138384	82.0312	0.0312500	990.4218
518400	1152.42	0.1093750	4997.848
2073600	18907.4	0.4843750	–
8433216	–	2.0937500	–
18045504	–	4.765625	–

MLFMM multilevel fast multipole method

The computation complexity of Algorithm 1, which is apparent, is $O(N^2)$. The computation time for Algorithm 1 is much less compared with the later MLFMM computation time when the unknowns are less than 100,000; however, the computation time for interior edge finding is comparatively long when the unknowns are extremely large, which is shown in Table 62.1, and all the problems are solved through notebook for Intel(R) Pentium(R) M 2.13 GHz CPU with 2 GB memory. Fast algorithm should be introduced to reduce the computing time for finding interior edges with millions of unknowns. The conventional fast method is to apply the Adelson-Velsky and Landis (AVL) tree with the computation complexity of $O(N\log(N))$ [5]. Although this method is fast and highly efficient, it is very complicated to write its code. In the next section, a new fast and simple algorithm is introduced in this chapter. The performance of this new algorithm is discussed in Sect. 3.

62.2 A New Fast and Simple Edge Finding

62.2.1 Finding Interior Edge for RWG Basis Function

The reason for Algorithm 1 with the computational complexity of $O(N^2)$ is to search the gotten interior edges one by one to find whether the new edge is resisted in the gotten interior edges, so the core of a fast algorithm is to search the computed interior edge quickly. The RWG basis function is used to illustrate the new algorithm. Assuming n is the total number of points in the mesh, a pointer array named node1 (Fortran language, for instance) with size n is introduced. Given the edge node list in natural number series, such as 1,2,3, …, n, the pointer array subscript maps to the first edge node list. The array pointer points to a two-dimensional array named node 2 having shape (m, 2). Given a node numbered with k, connected with m edges, where k is also the pointer array indicator, Fortran language is adopted to express node1(k)%node2 (m, 2). The first column of the two-dimensional array is

used to store the second node number, and the second column is used to store the global interior edge number. This pointer array forms a search data structure, which is shown in Fig. 62.3. The size of the pointer array equals to the number of points in the mesh, and the index of array is also the node number. One-dimensional array Node1Edge having shape (n) is used to store the total number of edges of each connected node. First, all triangular is scanned, and the sequence of points is rearranged according to the node number. Then, the number of edges associated with each node is counted and stored in an one-dimensional array Node1Edge, and then the memory of pointer array points is allocated dynamically, and Algorithm 2 is as follows:

Algorithm 2:

```
1  ne=0
2  Node1Edge(1:TotalNode)=0
3  Foreach triangular
4    Foreach edge of triangular nodes named no1,no2
5    flag = 0
6    For k1=1:Node1Edge(no1)
7      If node1(no1)%node2(k1,1)=no2 then
8        k= node1(no1)%node2(k1,2)
9        flag = 1
10        go to 20
11      end if
12    End
13    if  flag =0  then
14      ne = ne + 1
15      Node1Edge (no1)= Node1Edge (no1)+1
16      node1(no1)%node2(Node1Edge (no1),1)=no2
17      node1(no1)%node2(Node1Edge (no1),2)=ne
18      add new edge to the edge list
19    end if
20  End
21 End
```

The registered interior edge can be searched quickly from Steps 6–12 in Algorithm 2 through the searching data structure for the average of Node1Edge(:) which is a small constant. Algorithm 2 features fast speed because of its direct access to the

Node #	Connected Nodes	Connected Edges
1	2	1
	3	2
2	3	3
	4	4
3	4	5
	5	6
4	5	7
	6	8
5	6	9
6	-	-

Fig. 62.3 Searching data structure: node and edge connectivity lists

first node and the small constant number of average edges connected to each node. Table 62.1 shows that the computation time is comparatively less and increases linearly with the unknowns. The CPU time for computing interior edges is only 4.77 s with 18045504 unknowns for the new algorithm and increases linearly with the number of unknowns. The dash ("–") means the computational time is too long to compute interior edges, or the problem is too large to be computed by MLFMM on our notebook test platform.

62.2.2 Finding Interior Face for Schaubert–Wilton–Glisson Basis Function

The algorithm for face finding with Schaubert–Wilton–Glisson (SWG) basis function, used to discretize the volume integral equation, is similar to that of an RWG basis function [6]. A pointer array named node1 is applied to construct the searching data structure with its size equivalent to the number of tetrahedron nodes and points to a two-dimensional array named node2 with shape (k, 3), where k is the number of the facet which connected a node. The first column of two-dimensional array is used to store the second node index of the facet, where the SWG basis function is defined, the second column is used to store the third node index of the face, and the third column is used to store the global index of face. The number of rows of the two-dimensional array, the pointer array pointed, is determined by the number of faces which connected each node. All the tetrahedrons are scanned first, then also the total faces, so that each associated node can be gotten; in this sense, the memory can be allocated dynamically. The detail of the algorithm is analogous to algorithm 2 and is omitted here. Table 62.2 shows that the computation time for the new algorithm is comparatively less and increases linearly with the number of tetrahedrons.

Table 62.2 Time for computing interior faces and MLFMM (dielectric sphere)

Unknowns	Algorithm 1(s)	Algorithm 2(s)	MLFMM(s)
20777	3.000	0.1875	1813.156
95633	64.12500	0.6250	6063.625
208755	302.6406	1.1093	11744.88
425686	1258.86	2.1718	23134.93
1308296	11886.3	6.1093	-
3105696	-	14.35938	-

MLFMM multilevel fast multipole method

62.3 Performance of the New Algorithm

A high-quality mesh ought to have as many equilateral triangles as possible for the surface mesh with triangular cells. Figure 62.1 shows that there are n triangulars connected with point p. The angle of each triangular associated with point p is $\alpha_1, \alpha_2, \alpha_3 \cdots, \alpha_n$.

$$\alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_n = 360^\circ \tag{62.1}$$

As to the high quality of the mesh $\alpha_1, \alpha_2, \alpha_3 \cdots, \alpha_n \to 60^\circ$, we can get $n \to 6$; thus, the average number of edges associated with each node should be 6 or a constant value without changing apparently with the mesh size. Numerical tests are shown in Table 62.3, and the average edge is 6, which conforms to the analysis. More numerical experiments on simple geometry, such as hexahedron, cylinder, prism, and complex geometry including airplane, missile, worship, and so on, show that the average edge number is 6. In this sense, the computational complexity of the new algorithm is $O(N)$, and the memory requirement is $O(N)$ for the surface mesh Fig. 62.4.

The high-quality mesh ought to have as many equilateral tetrahedrons as possible for the volume mesh with tetrahedron cells; therefore, the average number of faces associated with each node should be a constant. A numerical experiment

Table 62.3 Average edges (surface mesh) or faces (volume mesh) associated with the edge node

RWG		SWG	
Unknowns	Average edges	Unknowns	Average faces
11664	6.0	23268	19.1
20736	6.0	52679	20.2
46656	6.0	88909	20.8
186624	6.0	169480	21.3
331776	6.0	389844	22.0
518400	6.0	1308296	22.6

RWG Rao–Wilton–Glisson, *SWG* Schaubert–Wilton–Glisson

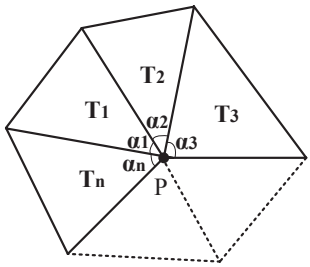


Fig. 62.4 Triangular mesh connected with point *p*

Table 62.4 Average faces associated with a node

Cylinder		Cube		Cone	
Unknowns	Average faces	Unknowns	Average faces	Unknowns	Average faces
8170	16.9	34208	17.2	10442	18.0
37304	19.3	148842	19.0	48182	20.2
124186	20.7	500070	20.4	161480	21.3
505956	21.8	1175259	21.1	1295442	22.5
2311764	22.6	3972236	21.9	4679534	23.1

is shown in Table III, and the average face connected with each node for dielectric sphere varies from 19.1 to 22.6; the number is increasing with the growing of the mesh size. More numerical experiments on cylinder, cube, and cone in Table 62.4 indicate that the change trend is similar to that of a sphere. So the computational complexity of the new algorithm is $O(N)$, and the memory requirement is $O(N)$ as to the volume mesh.

62.4 Conclusion

The new algorithm is simple and fast, which satisfies the demand of large-scale electromagnetic computational problems.

References

1. Rao SM, Wilton DR, Glisson AW. Electromagnetic scattering by surfaces of arbitrary shape. IEEE Trans Antennas Propag. 1982;AP-30(3):409–18.
2. Chew C, Jin JM, Midielssen E, Song JM. Fast and efficient algorithms in computational electromagnetics. Boston: Artech House Publishers; 2001. p. 56–249.
3. Hastriter ML. A study of MLFMA for large-scale scattering problems. PH. D. dissertation, Illinois, USA: Univ. Illinois at Urbana-Champaign; 2003.
4. Ergül Ö, Gürel L. Efficient parallelization of the multilevel fast multipole algorithm for the solution of large-scale scattering problems. IEEE Trans Antennas Propag. 2008;AP-56(8):2335–45.

5. Weiss MA. Algorithms data structures and problem solving with c++. Boston: Addison-Wesley Publishing Company; 1996. p. 20–217.
6. Schaubert DH, Wilton DR, Glisson AW. A tetrahedral modelling method for electromagnetic scattering by arbitrarily shaped inhomogeneous dielectric bodies. IEEE Trans Antennas Propag. 1984;AP-32(1):77–85.

Chapter 63
Test Point Selection Method Research Based on Genetic Algorithm and Binary Particle Swarm Optimization Algorithm

Xiaofeng Lv, Ling Ma, Jing Sun and Wei Pang

Abstract Test point selection is the foundation of testability analysis and design. A minimal complete subset of genetic algorithm and binary particle swarm optimization algorithm is proposed to meet testability index requirements. Firstly, the mathematical model is established based on analyzing the testability problems. Then, the heuristic function is constructed to measure the pros and cons of the test set. Experimental results show that the algorithm can effectively overcome the deficiency of a single algorithm going into a local optimum and premature convergence, and improve the searching efficiency to obtain a global optimal solution quickly.

Keywords Test point selection · GABPSO algorithm · Heuristic function

63.1 Introduction

With the increase of performance and complexity of weapon systems, it is more and more difficult to carry out the fault detection and diagnosis; thus, the development of testability analysis and design has become a pressing matter of the moment [1]. The purpose of test selection is to find out the optimal test point combination under the condition of meeting the system fault detection rate (FDR) and the fault isolation rate (FIR) among all test points so as to make the test cost (including all kinds of test cost and artificial cost) as small as possible. From the point of view of

X. Lv (✉) · L. Ma · J. Sun · W. Pang
Naval Aeronautical and Astronautical University, 264001 Yantai, China
e-mail: lxf_00112@163.com

L. Ma
e-mail: miller5356@vip.sina.com

J. Sun
e-mail: 19804473@qq.com

W. Pang
e-mail: 19804473@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_63

mathematics, the test point selection is a problem in terms of combinatorial optimization, which can be described by the set covering model; nevertheless, as the set covering is an N-P complete problem, when the system scale is large, it is usually difficult to gain an optimal solution. Currently, many literatures put forward corresponding solution methods [2–7], among which the genetic algorithm (GA) and the binary particle swarm optimization (BPSO) have reached a certain effect; however, the solving efficiency and accuracy are not ideal due to the inherent difficulty of the problem itself, and a new algorithm is demanded to obtain the optimal solution.

GA is a global optimization algorithm formed by referring to natural selection and group evolution extremum. The algorithm can effectively overcome the problem of falling into local minima in the optimization process, but its convergence is slow and efficiency is low [8]. The BPSO algorithm is an optimization algorithm based on population proposed by Kennedy and Eberhart in 1997. The BPSO algorithm features the advantages of simple structure and easy implementation; besides, it can gain a stable solution in short time, while the global searching ability is weak [9]. This chapter combines GA and BPSO to put forward a test point selection algorithm based on GABPSO. The chapter describes the test point selection problem and builds the mathematical model for test point selection first; then, considering the FDR and FIR, it defines the heuristic function of test points and gives the process of algorithm implementation; finally, the GABPSO algorithm can find the minimal complete subsets rapidly and effectively through examples.

63.2 Description of the Test Point Selection Problem

The test point selection problem indicates the selection of a suitable test subset among test points set under certain conditions so as to make the test subset meet the requirements of testing indices such as FDR and FIR, while requiring minimum test costs. If the test set T can meet the required testability design index, it is called the complete test set, excluding any test point in the given test set T which is no longer a complete test set and called the minimum complete test set. Finding the optimal test point set is to solve the minimum complete test set.

63.2.1 Fault-Test Dependency Matrix

Provided the fault set to be detected and isolated from the system is $F=\{F_i\}, i=1,2,\cdots,m$, the test point set available for selection is $T=\{t_j\}, j=1,2,\cdots,n$, and the fault set can be tested by the test points $F_j \subseteq F$. Similarly, as fault F_j can be detected by a series of test sets, the collection is $T_i \subseteq T$. The relationship between the test points and faults can be represented by the dependency matrix as follows:

$$FT = \begin{vmatrix} d_{11}d_{12}\cdots d_{1n} \\ d_{21}d_{22}\cdots d_{2n} \\ \vdots \quad \vdots \quad \vdots \\ d_{m1}d_{m2}\cdots d_{mn} \end{vmatrix} \tag{63.1}$$

Matrix FT is $m \times n$ Boolean matrix; the row of the matrix is the corresponding fault set, and the column is the corresponding test point set. In which d_{ij} is a Boolean variable, namely, if the fault F_i can be detected by testing point t_j, then $d_{ij} = 1$, otherwise $d_{ij} = 0$. According to the fault-test dependency matrix, some parameters which may reflect the fault detection and isolation capability of the alternative test set can be determined.

63.2.2 Test Selection Optimization Model

FDR is defined as the ratio of the number of faults detected correctly by a specified method and the total faults occurred of a unit under test within the specified time [1]. Assuming fault F_i can be detected by T_s, the row vector corresponding to F_i in the dependency matrix has at least one element 1, that is $\bigcup_{t_j \in T_S}^{N_S} d_{ij} = 1$. In the formula, $\cup$ is the OR operation of the Boolean variable and N_S is number of test point sets T_s. If the set consisting of all faults detected by T_s is F_D, then $F_D = \left\{ F_i \mid F_i \in F, \bigcup_{t_j \in T_S}^{N_S} d_{ij} = 1 \right\} \subseteq F$. Provided $|F|$ expresses the number of sets F, then the FDR can be expressed as follows:

$$\gamma_{FD} = \frac{|F_D|}{|F|} = \frac{|F_D|}{m} \tag{63.2}$$

The FIR can be defined as using a method to isolate the ratio for number of faults, and the fault number can be detected at the same time [1]. If the fault F_i and F_j can be isolated, i th row and j th dependency matrix must be different.

Assuming that the set consisting of test points can detect the fault F_i to be T_{F_i}, and the set consisting of test points can detect fault F_j to be T_{F_j}, then the conditions for fault F_i can be isolated as shown in $T_{F_i} \oplus T_{F_j} = 1, \forall F_i \in F, F_i \neq F_j$. In which $\oplus$ represents XOR operation of sets; in other words, when the two sets are not different, the result is 1; therefore, when the fault F_i can be isolated by the alternative test set T_s, there is $T_{F_i} \subseteq T_F$. Assuming the fault set which can be isolated by alternative test set is F_S, then $F_S = \left\{ F_j \mid F_j \in F, T_{F_i} \oplus T_{F_j} = 1, \forall F_j \in F, F_i \neq F_j \right\}$. Therefore, the FIR can be expressed as:

$$\gamma_{FI} = \frac{|F_S|}{|F|} \tag{63.3}$$

To sum up, under the premise of meeting FDR and FIR in advance, gain the minimum test cost as far as possible; therefore, the test point selection optimization problem can be expressed as follows:

$$\begin{cases} \min \sum_{T_i \in T} c_i \\ s.t \\ \gamma_{FD} \geq \gamma_{FD}^* \\ \gamma_{FI} \geq \lambda_{FI}^* \end{cases} \tag{63.4}$$

In Formula (4), c_i stands for the test cost of test points $t_i, \gamma_{FD}^* \; and \; \lambda_{FI}^*$ stands for FDR and FIR required by the system.

63.3 Test Point Selection Based on GABPSO Algorithm

63.3.1 Heuristic Function

The purpose of test point optimization selection is to find a test set meeting FDR and FIR predetermined by the system; thus choose the test set with minimum test cost to isolate all faults. As the test point selection belongs to the N-P complete problem, it is hardly to be realized by the enumeration method; thus, we measure pros and cons of test points by setting the heuristic value of a single test point to provide a new method for searching a test set rapidly.

The smaller a fault can be measured, the smaller the probability of the fault will be detected. Assuming the testability of fault F_i is expressed by $k(F_i)$, the fault set that can be detected by the test point t_i is $S(t_i)(S(t_i) \subseteq F)$; then the minimum testability of test point t_i is defined as $N(t_i) = min\{k(S(t_i)); t_i \in T\}$. Assuming the fault F_i can only be tested by test points t_i and t_j, the testability is 2; if t_i is excluded, t_j will be selected; otherwise the fault is not measurable. In this sense, the selection of a small coverage test is conductive to detect and isolate faults as well as the next step of test point selection. For the pros and cons of individual test t_j, it can be measured by the heuristic function as follows:

$$h(t_j) = a \frac{N(t_j)}{|t_j| P(t_j)} \tag{63.5}$$

In the formula, $N(t_j)$ expresses the minimum testability of test point t_j, $|t_j|$ stands for the number of faults which can be detected by test points t_j, $P(t_j)$ is the average probability of fault occurrence because the test point t_j can detected, a is the coefficient, $P(t_j)$ is constant, $h(t_j)$ is proportional to the minimum testability, and inversely proportional to $|t_j|$. Therefore, the smaller the heuristic function value is, the higher the test priority will become.

63.3.2 GABPSO Algorithm

GA and BPSO have much in common, both of which originated in human's research in biology, and use the fitness value to evaluate a system and proceed a search. GA has broad space search capability and ability of mutation, strong capability of global search, and can effectively overcome local optimal problems during the process of searching; thus, it is suitable for large-scale parallel calculation. But it features the problem of slow searching speed, poor population mobility, and lack of memory functions, and it cannot refer to historical information [8–10]. While the BPSO algorithm highlights the advantages of a simple structure, easiness of acquisition, and memory function, and it can converge to a stable solution in short time, it is easy to fall into a local optimum. This chapter combined GA and BPSO algorithm to propose a hybrid algorithm for test point optimization. It can not only ensure a better information exchange mechanism but also avoid falling into a local optimization problem so as to enhance the searching speed and improve the success rate of the optimal solution. The test point optimization algorithm flow chart based on the hybrid genetic binary particle swarm optimization is shown in Fig. 63.1.

Implementation steps are given below:

Step 1: Initialize the parameters, including the population size *psize,* the particle swarm inertial factor ω, the accelerated factors c_1 and c_2, the genetic crossover probability *ps,* the mutation probability *pc,* the maximum speed $V_{\max}$, and the number of iterations *maxgen* .

Step 2: Initialize the population according to the test selection problems, get the population $pop = \{x_{ij}\}_{psize \times d}$, in which d is the number of alternative test points; and calculate the fitness value according to formula (6). The bigger the fitness value is, the better the performance of particles will be, and the current individual optimal location gbest and group optimal location pbest can be obtained.

$$f(T_S) = (\gamma_{FD} \times \gamma_{FI})^2 \left\{ C \times H - \sum_{T_i \in T_S} c_i h(t_i) \right\} \tag{63.6}$$

In the formula, $C = \sum_{i=1}^{d} c_i, H = \max(h(t_i))$.

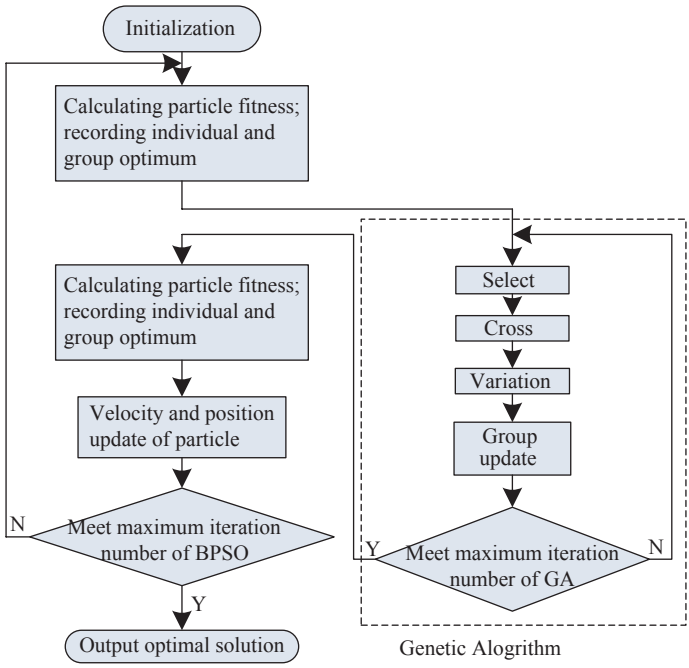


Fig. 63.1 Flow chart of GABPSO algorithm

Step 3: As to the probability of the individual value in proportion to the adaptive probability value, copy it to the new group and make the selection; do the cross operation with each of the two probabilities; then get the new group pop_1; do the mutation operation individually in the group; then get the new group pop_2.

Step 4: If GA satisfies the iteration number, then finish; calculate the fitness value for group pop_2 to select individual optimal position gbest and the best position of the swarm pbest; update the velocity and position of the particle to produce the next generation population; otherwise turn to step 3.

Step 5: If the maximum iterations are satisfied, the algorithm ends and outputs the optimal value of groups as a solution; otherwise turn to step 3.

63.4 Simulation and Verification

A certain type of missile-launching system fault-test dependency matrix is shown in Table 63.1. The system has a total of 18 faults and 22 alternative test points. The average probability of each fault is [0.01, 0.02, 0.05, 0.16, 0.03, 0.12, 0.28, 0.12, 0.11, 0.22, 0.28, 0.23, 0.09, 0.11, 0.24, 0.07, 0.19, 0.15], the cost of each test point is [1.0, 2, 1.3, 1.5, 1, 0.5, 3, 2.1, 2.2, 1.6, 1.7, 1.1, 1.5, 1.1, 1.4, 1.1, 0.9, 1.3, 1.5, 1.7, 0.9, 0.7]. As to the testability index of system, FDR is 100 %, and FIR is no less than 95 %.

Table 63.1 A certain type of missile launching system fault-test dependency matrix

Fault	Test points																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
2	1	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
4	0	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
6	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
8	0	0	1	1	0	1	0	1	1	1	0	0	1	0	0	0	0	0	1	1	1	0
9	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
12	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
13	0	0	1	0	0	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
14	0	0	1	0	0	0	0	1	0	1	1	1	1	0	1	0	0	1	0	1	1	1
15	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1
16	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1
17	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1
18	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1

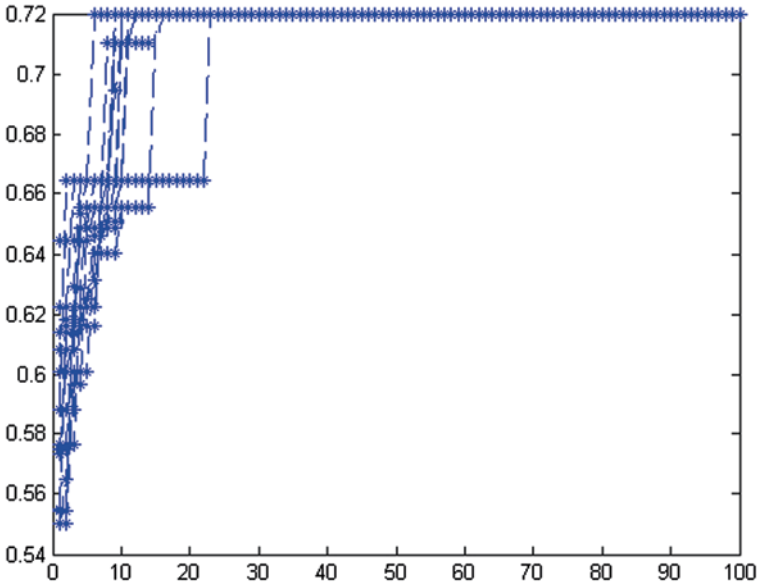


Fig. 63.2 Fitness curves of a certain type of missile launching system

Do the test point optimization by using the GABPSO algorithm, and set parameters value as follows: $psize = 20, maxgen = 100, c_1 = c_2 = 1.499, V_{\max} = 4$. After running for many times, the results obtained are [0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0], that is, the best test points set is $[t_3, t_4, t_5, t_8, t_{16}, t_{18}, t_{19}, t_{21}]$, at this time, the FDR of system is 100%, FIR is 100% and the test cost is 10.7. From Fig. 63.2, we find that GABPSO algorithm can obtain optimal solution in the 10th generation basically.

The simulation result shows that GABPSO algorithm proposed in this chapter has fast searching speed and high success rate, which can be effective to overcome the weaknesses of converging slowly and involving in local optimum.

63.5 Conclusion

With the increasing complexity of the equipment system, the relationship between fault and tests has become more and more complex. How to select the optimal test points set accurately and efficiently is becoming more and more important. GABPSO algorithm combines the advantages of GA and BPSO algorithm to effectively overcome the problem of premature convergence and local optimum in single algorithm through performance of complementary, thus enhancing the searching efficiency and improving the accuracy rate. Experimental results show that this method can find the globally optimal solution quickly and exactly under the premise of satisfying testability index of system, and it is of great significance to the testability analysis and design as to large complicated system.

References

1. Tian Z, Shi J. System testability design, analysis and verification. Beijing: Beijing University of Aeronautic and Astronautics Press; 2003. pp. 68–75.
2. Prasad VC, Rao Pinjala SN. Fast algorithms for selection of test nodes of an analog circuit using a generalized fault dictionary approach. Circuit Syst Signal Process. 1995;14(6):707–24.
3. Prasad VC, Babu NSC. Selection of test nodes of for analog fault diagnosis in dictionary approach. IEEE Trans Instrum Measur. 2000;49(6):1289–97.
4. Pinjala KK, Bruce CK. An approach for selection of test points for analog fault diagnosis. Proceedings of the 18th IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems, IEEE; 2003. pp. 287–94.
5. Starzyk JA, Liu D, Liu Z-H, et al. Entropy-based optimum test nodes selection for analog fault dictionary techniques. IEEE Trans Instrum Measur. 2004;53(3):754–61.
6. Kennedy J, Eberhart RC. A discrete binary version of the particle swarm algorithm. Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics, IEEE; 1997. pp. 4104–9.
7. Chenlin Y, Shulin T, Bing L. Application of heuristic graph search to test points selection for analog fault dictionary techniques. IEEE Trans Instrum Measur. 2009;58(7):2145–58.
8. Zhou M, Sun S. The theory and application of genetic algorithm. Beijing: National Defense of Industry Press; 1999. pp. 36–48.
9. Engelbrecht AP. Fundamentals of computational swarm intelligence. Hoboken: Wiley; 2009. pp. 46–64.
10. Ronghua J, Houjun W, Bing L. Test selection based on binary particle swarm optimization. J Electron Measur Instrum. 2008;22(2):11–5.

Chapter 64
Vehicle Weigh-in-motion Systems Based on Particle Swarm Optimization

Huijuan Ding, Quanhu Li, Ting Xu and Nengshao Li

Abstract In this chapter, the particle swarm optimization (PSO) algorithm based on the random global optimization is introduced into the network training in order to eliminate the problem in back-propagation (BP) neural networks of vehicle weigh-in-motion (WIM) systems which are sensitive to the initial weights, easy to fall into the local minimum, and have a slow convergence rate. We have established a PSO-BP neural network model to optimize the initial weighted threshold and structure of the neural network. Simulation results show that the PSO-BP neural network model has faster convergence rate and a higher precision.

Keywords Weigh-in-motion system · BP neural network · PSO algorithm

64.1 Introduction

In this chapter, we build an appropriate mathematical model of a dynamic weighing system according to the actual situation of Inner Mongolia provinces thoroughfare. The back-propagation (BP) neural network algorithm based on the gradient descent depends on the choice of initial weights, so it has the defects of slow convergence rate and easiness of falling into the local minimum [1]. The optimized BP neural network algorithm is based on particle swarm optimization (PSO) which maps the weights and threshold of networks to particles, and we then complete the network training through competition and cooperation between particles.

H. Ding (✉) · Q. Li · T. Xu · N. Li
Department of Electronic and Information Engineering,
Inner Mongolia University, 010021 Hohhot, China
e-mail: dinghuijuan88@163.com

Q. Li
e-mail: 1514178566@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_64

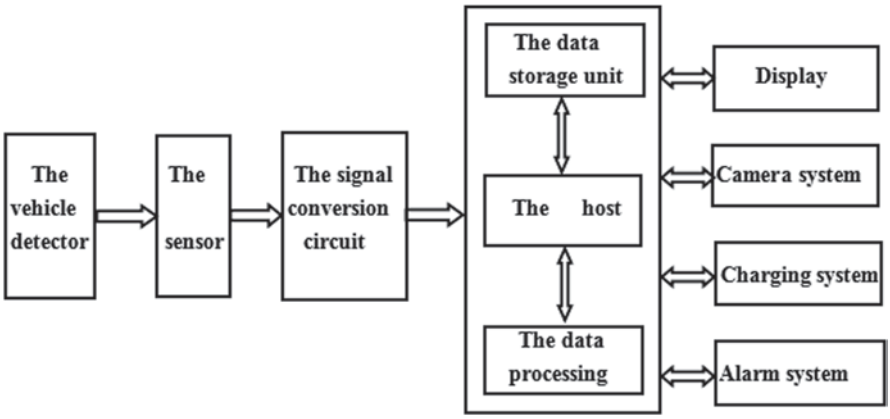


Fig. 64.1 Design diagram of the WIM system

64.2 Weigh-in-Motion System

Generally speaking, the vehicle weigh-in-motion (WIM) system mainly consists of the vehicle detector, the signal conversion circuit (including the data amplifier circuit, filter and analog-to-digital (A/D) converter), the host, the display, the charging system, and the alarm system. The composition of the vehicle WIM system is shown in Fig. 64.1.

The vehicle detector will upload the type and speed of the vehicle to the host when it enters the detection lane. The weight of the signal will be converted into a voltage signal by the pressure sensor. By amplifying, filtering and A/D conversion, the final signal will be sent to the host. Then, the host would build a vehicle dynamic weighting model according to the received signal data, and the model estimates the weight of the vehicle by the proposed algorithm and finally compares it with the actual weight to estimate whether or not the vehicle is overweighed. In addition, the camera system will take a photograph of the overweighed vehicles with the images and weighing data of the overweighed vehicles uploaded to the host [2]. The drivers can see the test results in the display in front of the toll station.

Although the filter in the signal conversion circuit is capable of removing the high-frequency noise in the data, the original signal pretreated though the signal conversion circuit is mixed with dynamic load caused by vehicle vibration, which cannot be filtered by the filter circuit because its frequency is too low. Generally, the frequency of the dynamic load ranges from 3 Hz to 20 Hz, with the amplitude of dynamic load equal to 10 % of the vehicle's weight [3]. The accuracy of dynamic weighing results cannot reach the demand without the dynamic load filtered. In this chapter, we propose a new algorithm for removing the dynamic load.

64.3 Data Processing Algorithm of the Vehicle WIM System

64.3.1 Particle Swarm Optimization

Assume that there are N particles in an H-dimensional target searching space. The coordinate position of each particle in the H-dimensional space is $x_i = (x_{i1}, x_{i1}, \cdots, x_{ih})$, and the velocity of each particle is $v_i = (v_{i1}, v_{i2}, \cdots, v_{ih})$. Particles update their velocity and position by the following formulas during each iteration:

$$v_{ih}^{k+1} = \omega v_{ih}^{k} + c_1 r_1 (p_{ih} - x_{ih}^{k}) + c_2 r_2 (p_{gh} - x_{ih}^{k}) \tag{64.1}$$

$$x_{ih}^{k+1} = x_{ih}^{k} + v_{ih}^{k+1}, \tag{64.2}$$

where $i = 1, 2, \ldots, N$; $h = 1, 2, \ldots, H$; K is the number of iterations; ω is the inertial weight; the value of r_1 and r_2 is a random number between [0, 1]; c_1 and c_2 are the learning factors [4].

The formula shows that the update of the particle's velocity includes three parts. The first part is the previous velocity of the particles. Adjustment of ω dynamically can balance the global and local search ability in the process of searching to ensure the particle's convergence toward the global optimal solution. The second part is to learn from the optimal position of the particles that can make the particles to have strong local search ability. The third part is to learn from the optimal position of the particles. The particles have the ability to learn from itself and excellent individuals in the group by adjusting the learning factors, so it can reach the individual extremum and global optimal value of particles [5]. The velocity of the particles in each dimension should be less than the maximum velocity set by algorithm in the process of iterative update.

64.3.2 Optimizing Neural Network Based on Particle Swarm

The neural network with nonlinear characteristic can deal with the uncertainty function between various factors in the dynamic weighing system of vehicles. In this chapter, the PSO algorithm is introduced into the network to improve the BP neural network. The optimization process of the PSO algorithm uses the interaction between particles rather than the gradient information. In addition to its easiness of operation, there are other advantages of the PSO algorithm because it can overcome the defect of the local minimum and improve the convergence rate of the local area [6].

64.3.2.1 Structure of PSO-BP Neural Network

The network adopts a three-layer network: the input layer, the hidden layer, and the output layer. The number of nodes in the input layer and the output layer is determined by specific situations. As the key of the weight accuracy is to solve the relationship between the weight and the velocity, we choose the weight and velocity of the vehicle which mainly influence the dynamic weighing as the input parameters in this chapter. As the purpose of the WIM system is to measure the real weight of the vehicle, the output parameter is the real weight of the vehicle. Since the inconsistency is available in the collected data units, in consideration of the difference between the magnitude and the variation of different parameters, the small range or absolute value of data may be concealed by other factors which will affect the network training effects; thus, the input and output data must be normalized into interval [−1, 1] before training the network. Upon the normalization of the data, the program is allowed to accelerate convergence at run time; finally, the data must be antinormalized.

The number of hidden layer nodes has great influences on the training precision and the convergence rate of the network. Preliminarily, we determine that the number of hidden layer nodes is 3–12 according to the empirical formula; then, we determine that the number of hidden layer nodes is 5 by trial and error; thus, the structure of the network is 2-5-1, which means the network has 2*5 + 5*1 + 5 + 1 total weight and threshold value, then the dimension of the particles would be 21 based on the dimension of the particle; therefore, there is a mapping between the dimensions of the particle and the connection values of the neural network.

We set the hidden layer transfer function as tansig, the output layer transfer function as purelin, the target accuracy of network training as 0.001, and the number of network training as 2000 times. Parameters of the PSO algorithms are set as follows: the total particle number is 15, c1 = c2 = 2.05, Wmax = 0.9, Wmin = 0.4, Vmax = 1.

64.3.2.2 Steps of the Optimizing BP Algorithm based on Particle Swarm

- Initialize the correlation parameters of the algorithm: the weights and threshold of network, the initial velocity and position of the particles, inertial weight, and the maximum number of iterations.
- The mean square error (MSE) of the BP neural network is selected to represent the particle swarm fitness function, so the formula is expressed as follows:

$$fitness = \frac{1}{m}\sum_{i=1}^{m}\sum_{k=1}^{n}(y_i^k - t_i^k)^2, \tag{64.3}$$

where y is the actual output of the output sample, t is the expected output of the output sample, and m is the number of the training sample set.

- The velocity and the position of every iteration are updated based on the formula of velocity and position; at the same time, we also have to consider whether the updated velocity and position are in a limited range.
- In the process of each iteration, we calculate the fitness value of the particle according to formula 3. Then we compare it with the individual extremum. If it is superior to the current individual extremum, we use it as the individual extremum. The global optimal value should also be updated when the best individual extremum is superior to the current global optimal value.
- Put the global optimal value of the last iteration as BP weights and thresholds of neural network for training. The network will put out the results of the simulation if they meet the preset accuracy or achieve the max iterations; otherwise, continue the training.

64.4 System Simulation and Analysis of the Results

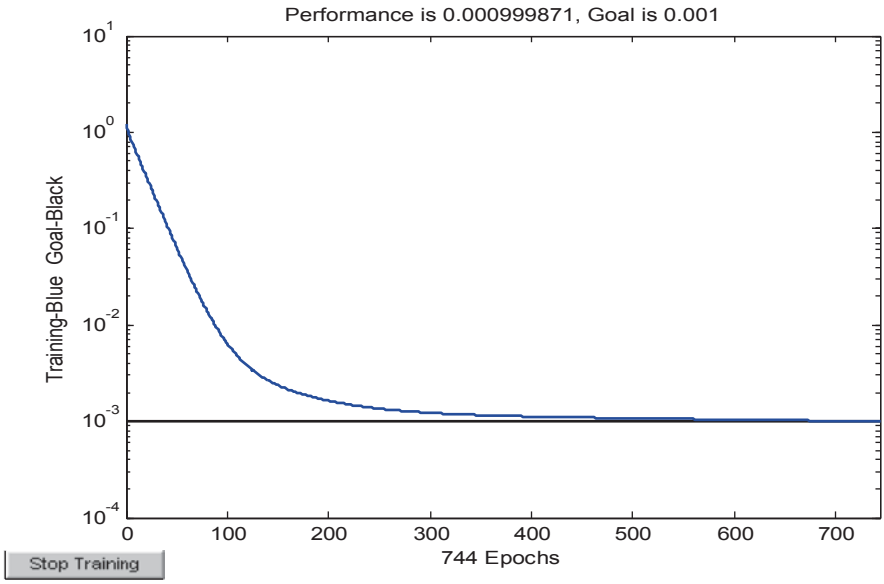
Sources of experimental data come from two axes trucks whose weights are 5000 and 15,000 kg, respectively, at the speed of 10–30 km/h obtained by the weighing sensor [7]. The data are divided into two parts, with one part of 200 groups of data used for the network training and another part of 68 groups of data for testing the trained network.

With the BP and PSO-BP network for the sample data training and testing, respectively, BP network training requires 744 steps to converge to the default precision while the PSO-BP network training can converge to the default precision just in 19 steps, as seen in Fig. 64.2a and b. In the process of training, 200 groups of data are used for network testing. Upon training the network, 68 groups of data are used to test the model. In the process of testing, only a small part of the BP network fitting value coincides with the actual value in Fig. 64.3; however, Fig. 64.4 shows almost all the fitting values, and actual values are consistent in the PSO-BP network.

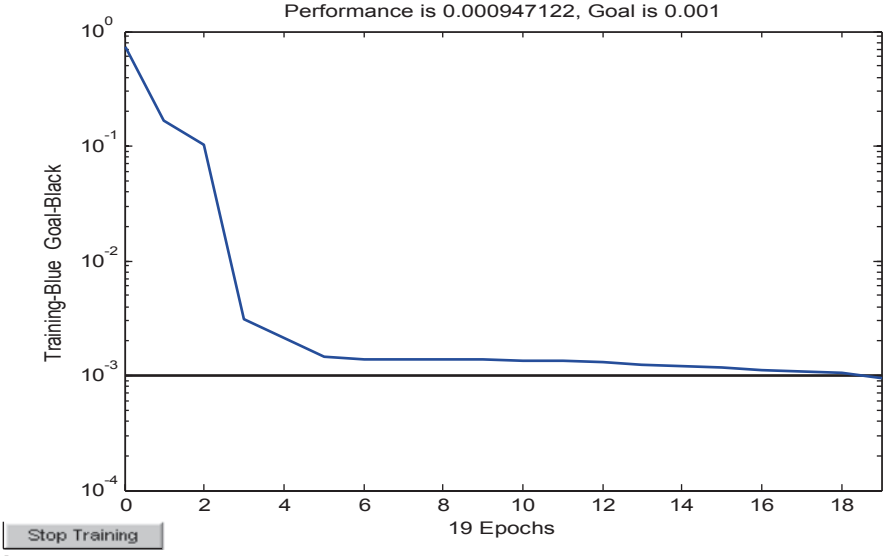
The test uses simulation to get the fitting values of the network and compares them with the actual values. In this way, we can judge whether the relative error between the actual values and fitting values meets the requirements. The relative error is calculated as follows:

$$Error = \frac{|W - T|}{T} * 100\%. \tag{64.4}$$

The relative error of 30 samples in the testing data is shown in Table 64.1. In the BP network testing, the smallest relative error in Table 64.1 is 1.3 %, and the average relative error of BP network is 2.56 %; however, in the PSO-BP network testing, the biggest relative error in Table 64.1 is 2.4 %; in this sense, it seems the performance of optimized BP algorithm based on PSO is better than the basic BP algorithm.



a



b

Fig. 64.2 **a** Convergence figure for BP. **b** Convergence figure for PSO

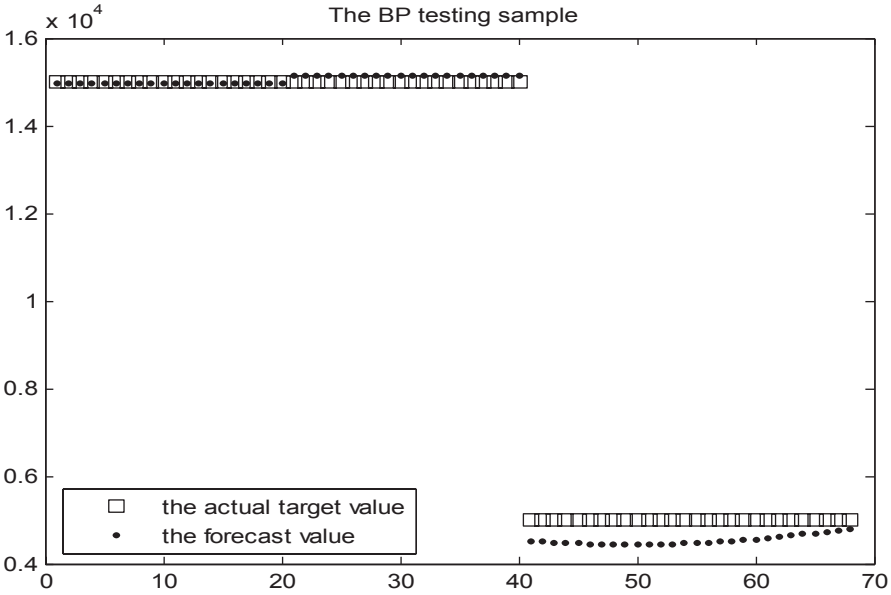


Fig. 64.3 Testing result for BP

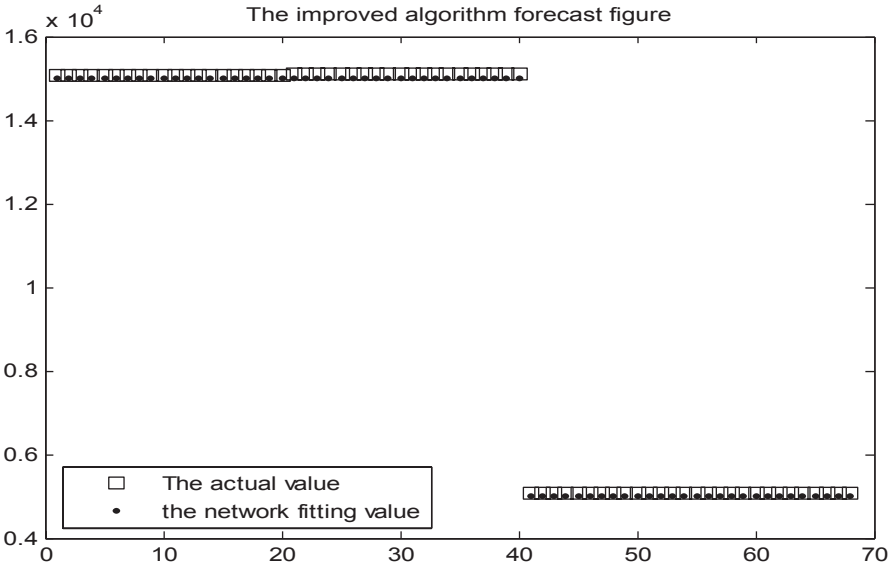


Fig. 64.4 Testing result of PSO

Table 64.1 Table of error analysis of test results

Sample	Velocity (km/h)	The real weight (kg)	The relative error of BP (%)	The relative error of PSO-BP (%)
1	5	15,000	1.9	1.7
2	5	15,000	1.7	1.2
3	5	15,000	1.3	1.1
4	5	15,000	1.5	0.9
5	5	15,000	1.4	0.7
6	5	15,000	1.5	1.2
7	5	15,000	1.7	1.4
8	5	15,000	1.6	1.3
9	5	15,000	1.5	1.2
10	5	15,000	1.9	1.5
11	12	15,000	2.6	1.6
12	12	15,000	2.9	1.6
13	12	15,000	2.5	1.1
14	12	15,000	2.7	1.2
15	12	15,000	2.2	1.4
16	12	15,000	2.5	1.7
17	12	15,000	2.3	1.5
18	12	15,000	2.1	1.4
19	12	15,000	2.5	1.1
20	12	15,000	2.4	1.3
21	16	5000	3.5	2.0
22	16	5000	3.1	2.1
23	16	5000	3.4	2.4
24	16	5000	3.9	1.9
25	16	5000	3.5	1.7
26	16	5000	4.1	1.6
27	16	5000	4.3	1.9
28	16	5000	3.9	2.1
29	16	5000	3.2	2.0
30	16	5000	3.1	2.3

64.5 Conclusion

In this chapter, we train the sample data with BP algorithm first, which has used the gradient descent information to adjust its connection weights, but it is easy to fall into the local optimum value. Thus, the PSO with global search capability has been introduced to optimize the BP algorithm and the PSO-BP algorithm model is thus

proposed. The simulation results show that the PSO-BP algorithm of the dynamic vehicle weighing model has significantly reduced the number of iterations and features high prediction accuracy.

References

1. Li Li Niu B. Particle swarm optimization algorithm. Beijing: Metallurgical industry press; 2009. p. 108–11. (In Chinese)
2. Chang Q, Zhang Y, Yu Z. Application of BP neural network to dynamic weighing with cars. J Electron Qual. 2011;4(5):37–8. (In Chinese)
3. Ling J. Study on the design theory of highway weigh-in-motion system. Chang ’an: Chang ’an University; 2001. (In Chinese)
4. Liang Y, Wu C, Shi X, Ge H. Theory and application of swarm intelligence optimization algorithm. Beijing: Science Press; 2009. p. 76–8. (In Chinese)
5. Liu Y, Niu B. Theory and application of a new particle swarm optimization. Beijing: Science press; 2013. p. 15–8. (In Chinese)
6. Ai Y. The modeling and application in stock market forecast based on neural network with particle swarm optimization. Hefei: Hefei University of Technology; 2009. (In Chinese)
7. Yang M. Research on the algorithm of vehicle weigh-in-motion system. Chang ’an: Chang ’an University; 2010. (In Chinese)

Chapter 65
An Improved Unscented Kalman Filter Algorithm Based on QR Decomposition

Dongmei Li, Chengdong Xu and Xi Cai

Abstract The principle of extended Kalman filter (EKF), unscented Kalman filter (UKF) algorithm, and unscented transformation (UT) is studied. To simply the process of obtaining sigma points in UKF, an improved UKF algorithm is mentioned in this chapter. The theory formulation and computer simulation have been done for the comparison of these three nonlinear Kalman filters. Results show that all of them can accomplish the estimation task, but the improved UKF has higher accuracy and less computation cost than the other two.

Keywords EK · UKF · QR decomposition

65.1 Introduction

Nonlinear filtering is developed after linear filtering; with the rapid development of the control and communications, nonlinear filtering techniques have become increasingly demanding [1]. Due to that in most cases, the system model is not necessarily linear and a Gaussian distribution; the optimal solution cannot be obtained by the classical Kalman filter. Thus, the study of the nonlinear filtering method becomes especially important. Extended Kalman filter (EKF) is the most common in the nonlinear filtering problems, which transforms the nonlinear filtering problems into a linear filtering problem by linearization. However, the error from linearization of the model tends to seriously affect the final filtering accuracy and even leads to filter divergence [2]. To solve this problem, the unscented Kalman filter (UKF) is proposed, which obtains sampling points by unscented transformation (UT) to approximate the posterior distribution of the state vector, without linearization of nonlinear systems [3]. Therefore, it can avoid the linearization error and make the state mean value and variance reach the result that a nonprecision linear system is expanded by second-order Taylor series.

D. Li (✉) · C. Xu · X. Cai
Key Laboratory of Dynamics and Control of Flight Vehicle, Ministry of Education, School of Aerospace Engineering, Beijing Institute of Technology, 100081 Beijing, China
e-mail: ldmworkhard@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_65

The rest of this chapter is organized as follows: The two nonlinear Kalman filter algorithms EKF and UKF are introduced in Sect. 65.2. An improved UKF algorithm is mentioned in Sect. 65.3. In Sect. 65.4, numerical simulations are studied for a nonlinear system model. At last, some conclusions are drawn in Sect. 65.5.

65.2 Nonlinear Kalman Filter

State and observation equation of nonlinear systems are as below:

$$\begin{cases} \boldsymbol{X}_{k+1} = f_k(\boldsymbol{X}_k) + \boldsymbol{W}_k \\ \boldsymbol{Z}_k = h_k(\boldsymbol{X}_k) + \boldsymbol{V}_k \end{cases}, \tag{65.1}$$

where $\boldsymbol{X}_k$ denotes the state vector, and $\boldsymbol{Z}_k$ denotes the observation vector. $\boldsymbol{W}_k$ and $\boldsymbol{V}_k$, respectively, denote the process noise with n-dimension and the measurement noise with m-dimension.

65.2.1 Extended Kalman Filter

Let $\hat{\boldsymbol{X}}_{k,k}$ be the state estimate at time k, and $\boldsymbol{P}_{k,k}$ be the estimation error covariance at the same time [4–6]. Figure 65.1 illustrates the algorithm flow of the EKF.

In Fig. 65.1, $\boldsymbol{K}_{k+1}$ is the Kalman gain; $\boldsymbol{Q}_k$ is the variance of the process noise; and $\boldsymbol{F}_k = \left[\frac{\partial f_k}{\partial X_k}(\hat{\boldsymbol{X}}, \boldsymbol{W}_k)\right]$, $\boldsymbol{G}_k = \left[\frac{\partial f_k}{\partial W_k}(\hat{\boldsymbol{X}}_k, \boldsymbol{W}_k)\right]$ and $\boldsymbol{H}_{k+1} = \left[\frac{\partial h_{k+1}}{\partial X_k}(\hat{\boldsymbol{X}}_{k+1,k}, \boldsymbol{V}_{k+1})\right]$ are the Jacobins of the system equation and the measurement equation.

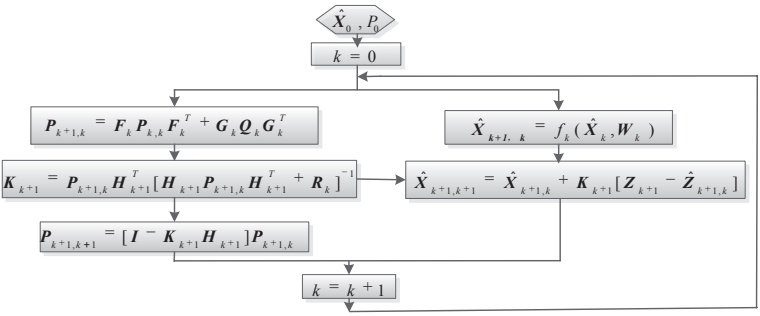


Fig. 65.1. The algorithm flow chart of the EKF

65.2.2 Unscented Kalman Filter

Because the EKF only uses the first-order terms of the Taylor series expansion of the nonlinear functions, it often introduces large errors in the estimated statistics of the posterior distributions of the states. This is especially evident in the way when the models are highly nonlinear and the local linearity assumption breaks down. Unlike the EKF, the UKF does not approximate the nonlinear process and observation models, but the UKF takes UT principles combined with the Kalman filter and then obtains the mean and covariance of the posterior distribution by the "sigma points" [7–9].

65.2.2.1 Unscented Transformation

For any nonlinear function $\boldsymbol{Z} = g(x)$, $\bar{\boldsymbol{x}}$ and $\boldsymbol{P}_x$ are, respectively, mean value and variance of the $\boldsymbol{x}$. To calculate the mean and the covariance matrix of the $\boldsymbol{Z}$, the UT is considered [10–11]. It approximately describes the random variable $\boldsymbol{x}$ by a set of deterministic chosen points, the so-called sigma points with corresponding weights.

$$\begin{aligned}
&\boldsymbol{\chi}_0 = \bar{\boldsymbol{x}} && w_0^m = \lambda/(n+\lambda) && i = 0 \\
&\boldsymbol{\chi}_i = \bar{\boldsymbol{x}} + (\sqrt{(n+\lambda)\boldsymbol{P}_x})_i && \lambda/(n+\lambda) + (1-\alpha^2+\beta) && i = 1,2,\cdots,n \\
&\boldsymbol{\chi}_i = \bar{\boldsymbol{x}} - (\sqrt{(n+\lambda)\boldsymbol{P}_x})_i && w_i^m = w_i^c = 1/2(n+\lambda) && i = n+1, n+2,\cdots,2n
\end{aligned} \tag{65.2}$$

where $\lambda = \alpha^2(n+\kappa) - n$ is a scaling parameter with the constant parameter, and $0 \le \alpha \le 1$, β, and κ are the tuning parameters. w_i^m and w_i^c are the weights of corresponding points. Each sigma point is now propagated through the nonlinear function $\boldsymbol{Z}_i = g(\boldsymbol{\chi}_i)$, and the estimated value and covariance can be computed as follows:

$$\bar{\boldsymbol{Z}} = \sum_{i=0}^{2n} w_i^m \boldsymbol{Z}_i, \boldsymbol{P}_{\mathrm{Z}} = \sum_{i=0}^{2n} w_i^c (\boldsymbol{Z}_i - \bar{\boldsymbol{Z}})(\boldsymbol{Z}_i - \bar{\boldsymbol{Z}})^{\boldsymbol{T}}. \tag{65.3}$$

65.2.2.2 UKF Algorithm Flow

The UKF works in such a way that at discrete sampling time k, $2n+1$ sigma points are generated; then the transformed sigma points are calculated by instantiating each point through a nonlinear function, and the weighted mean and covariance of the transformed points are calculated. After that, the state and covariance are updated for the next calculation. Figure 65.2 illustrates the algorithm flow of the UKF.

We can see that the UKF can estimate the mean and covariance of the state with a second order for any nonlinear system, and the Jacobians are not needed.

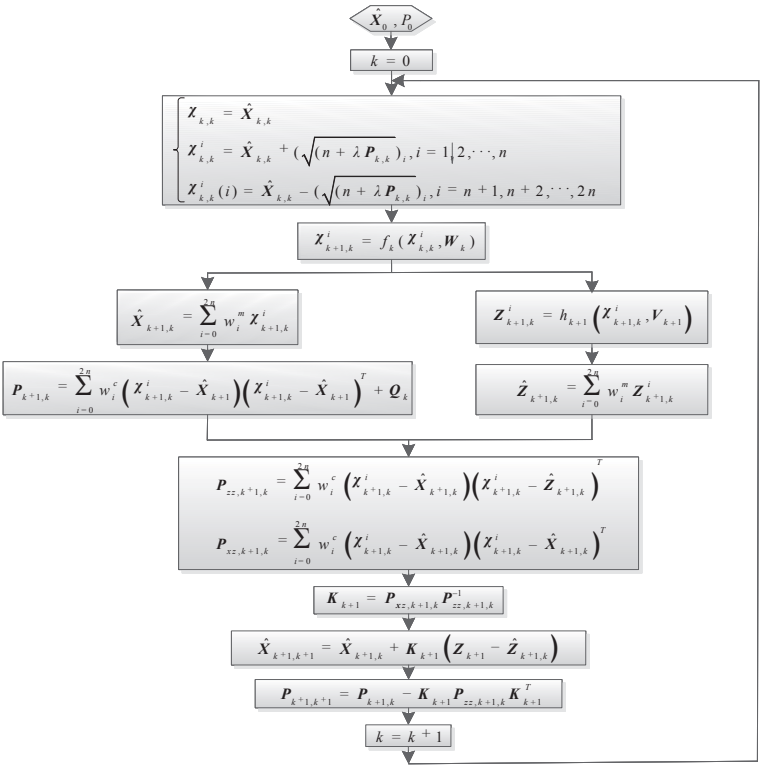


Fig. 65.2 The algorithm flow chart of the UKF

65.3 Improved UKF Algorithm

In the UKF algorithm, the new sigma point must be calculated each updating time, and the square root of the state covariance matrix P is needed to recalculate. Without doubt it will increase the amount of calculation and bring a corresponding estimation error. In this chapter, based on the original algorithm, an improved UKF algorithm is given, which used QR decomposition and Cholesky factor updating to update the square root of the covariance matrix P [12]. Set $\boldsymbol{SS}^T = \boldsymbol{P}$; in the improved UKF, $\boldsymbol{S}$ will be propagated directly, thus avoiding the need to recalculate at each time step.

- QR decomposition: For a matrix $\boldsymbol{A} \in R^{L\times N}$, $\boldsymbol{A}^T = \boldsymbol{QR}$ (Q is an orthogonal matrix $\boldsymbol{Q} \in \boldsymbol{R}^{N\times N}$ and $\boldsymbol{R}$ is an up-triangle matrix $\boldsymbol{R} \in \boldsymbol{R}^{N\times L}$). The up-triangle part $\tilde{\boldsymbol{R}}$ is the transposition of the Cholesky decomposition factor $\boldsymbol{P} = \boldsymbol{AA}^T$, $\tilde{\boldsymbol{R}}^T \tilde{\boldsymbol{R}} = \boldsymbol{AA}^T$. Let $qr\{\bullet\}$ denote the QR decomposition.
- Cholesky factor updating: If $\boldsymbol{S}$ is the initial Cholesky factor of $\boldsymbol{P}$, the Cholesky factor of $\boldsymbol{P} \pm \sqrt{v}\boldsymbol{uu}^T$ is denoted as $\boldsymbol{S} = \text{cholupdate}\{\boldsymbol{S}, u \pm v\}$.

- The improved UKF is shown as follows:

1. Initialization

$$\begin{cases} \hat{\boldsymbol{X}}_0 = E[\boldsymbol{X}_0] \\ \boldsymbol{S}_0 = E[(\boldsymbol{X}_0 - \hat{\boldsymbol{X}}_0)(\boldsymbol{X}_0 - \hat{\boldsymbol{X}}_0)^T] \end{cases} \tag{65.4}$$

2. Calculation of the sample point

$$\begin{cases} \boldsymbol{\chi}_{k,k}^i = \hat{\boldsymbol{X}}_{k,k} + \left(\sqrt{(n+\lambda)}\left(\boldsymbol{S}_{k,k}\right)_i, i = 1,2,\ldots,n \right. \\ \boldsymbol{\chi}_{k,k}^i(i) = \hat{\boldsymbol{X}}_{k,k} - \left(\sqrt{(n+\lambda)}\left(\boldsymbol{S}_{k,k}\right)_i, i = n+1, n+2,\ldots 2n \right. \end{cases} \tag{65.5}$$

3. Time updating

$$\hat{\boldsymbol{X}}_{k+1,k} = \sum_{i=0}^{2n} w_i^m \boldsymbol{\chi}_{k+1,k}^i \tag{65.6}$$

$$\boldsymbol{S}_{k+1,k} = qr\left\{[\sqrt{w_1}(\hat{\boldsymbol{X}}_{1:2n,k+1,k} - \hat{\boldsymbol{X}}_{k+1,k})\sqrt{\boldsymbol{Q}_{k-1}}]\right\} \tag{65.7}$$

$$\boldsymbol{S}_{k+1,k} = \text{cholupdate}\left\{\boldsymbol{S}_{k+1,k}, (\hat{\boldsymbol{X}}_{0,k+1,k} - \hat{\boldsymbol{X}}_{k+1,k}), w_0\right\} \tag{65.8}$$

$$\boldsymbol{Z}_{k+1,k}^i = h_{k+1}(\boldsymbol{\chi}_{k+1,k}^i, \boldsymbol{V}_{k+1}) \tag{65.9}$$

$$\hat{\boldsymbol{Z}}_{k+1,k} = \sum_{i=0}^{2n} w_i^m \boldsymbol{Z}_{k+1,k}^i \tag{65.10}$$

4. Measurement updating

$$\boldsymbol{S}_{ZZ,k+1,k} = qr\left\{[\sqrt{w_1}(\hat{\boldsymbol{Z}}_{1:2n,k+1,k} - \hat{\boldsymbol{Z}}_{k+1,k})\sqrt{\boldsymbol{Q}_k}]\right\} \tag{65.11}$$

$$\boldsymbol{S}_{ZZ,k+1,k} = \text{cholupdate}\left\{\boldsymbol{S}_{ZZ,k+1,k}(\hat{\boldsymbol{Z}}_{0,k+1,k} - \hat{\boldsymbol{Z}}_{k+1,k}), w_0\right\} \tag{65.12}$$

$$\boldsymbol{K}_{k+1} = (\boldsymbol{P}_{xz,k+1,k} / \boldsymbol{S}_{ZZ,k+1,k}) / \boldsymbol{S}_{ZZ,k+1,k}^T \tag{65.13}$$

$$\hat{\boldsymbol{X}}_{k+1,k+1} = \hat{\boldsymbol{X}}_{k+1,k} + \boldsymbol{K}_{k+1}(\boldsymbol{Z}_{k+1} - \hat{\boldsymbol{Z}}_{k+1,k}) \tag{65.14}$$

$$\boldsymbol{P}_{xz,k+1,k} = \sum_{i=0}^{2n} w_i^c (\boldsymbol{\chi}_{k+1,k}^i - \hat{\boldsymbol{X}}_{k+1,k})(\boldsymbol{\chi}_{k+1,k}^i - \hat{\boldsymbol{X}}_{k+1,k})^T \tag{65.15}$$

$$\boldsymbol{U} = \boldsymbol{K}_{k+1}\boldsymbol{S}_{zz,k+1,k} \tag{65.16}$$

$$\boldsymbol{S}_{k+1} = \text{cholupdate}\left\{\boldsymbol{S}_{k+1,k}, \boldsymbol{U}, -1\right\} \tag{65.17}$$

In Eq. 65.7, the time updating of the Cholesky factor $\boldsymbol{S}_{k+1,k}$ is calculated using QR decomposition of the compound matrix containing the weighted propagated sigma points, and in Eq. 65.8, the subsequent Cholesky updating is necessary since the zero weight. The same two steps are used to calculate the Cholesky factor $\boldsymbol{S}_{zz,k+1,k}$ of the observation error covariance in the Eqs. 65.11 and 65.12. The posterior measurement updating of the Cholesky factor about the state covariance is calculated in Eq. 65.17.

65.4 Simulation Analysis

To test the improved UKF algorithm and to compare with the existing algorithms EKF and UKF, a nonlinear system is used in this simulation. The dynamic state equation of it is expressed as below:

$$\begin{cases} x_n = \alpha x_{n-1} + \beta \dfrac{x_{n-1}}{1 + x_{n-1}^2} + \gamma \cos(1.2(n-1)) + w_n \\ y_n = \dfrac{x_n^2}{20} + v_n \end{cases} \tag{65.18}$$

where $w_n \sim N(0,1)$ and $v_n \sim N(0,1)$ are white Gaussian noise, and $\alpha = 0.5$, $\beta = 25$, and $\gamma = 8$ are simulation parameters. The system initialization parameters are set: initial state $x_0 = 0.1$ and initial variance $p_0 = 2$.

Figures 65.3 and 65.4 show the state estimate of the EKF, UKF, and improved UKF for this system. It is seen from the figures that the result of the improved UKF has a good agreement with the true value. The root mean square error (RMSE) is shown as Fig. 65.5, from which it can be seen that the RMSE of improved UKF is less than the others.

Table 65.1 gives the computation times and the mean value of the RMSE, in which the improved UKF needs a shorter computation time than the standard UKF, but the EKF has a shorter computation time than the standard UKF.

RMSE is used to measure the deviation between the observed value and the true value. T denotes the time step at one simulation [13].

$$\text{RMSE} = \left(\frac{1}{T} \sum_{k=1}^{T} (\hat{x}_k - x_k)^2 \right)^{1/2} \tag{65.19}$$

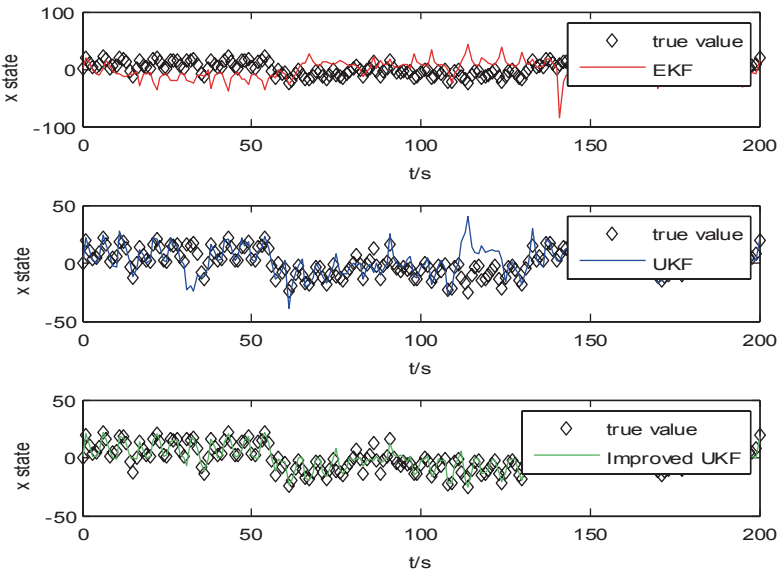


Fig. 65.3 State estimate. *EKF* extended Kalman filter, *UKF* unscented Kalman filter

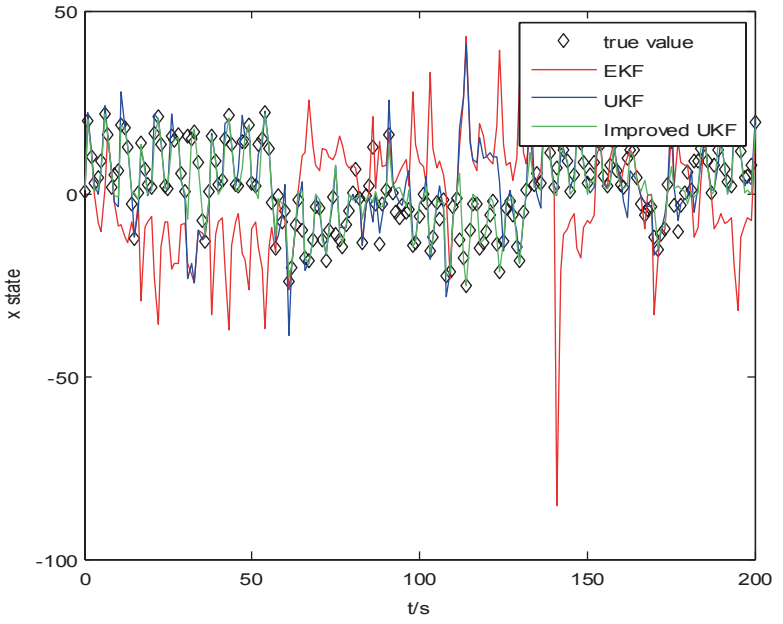


Fig. 65.4 State estimate of three algorithms. *EKF* extended Kalman filter, *UKF* unscented Kalman filter

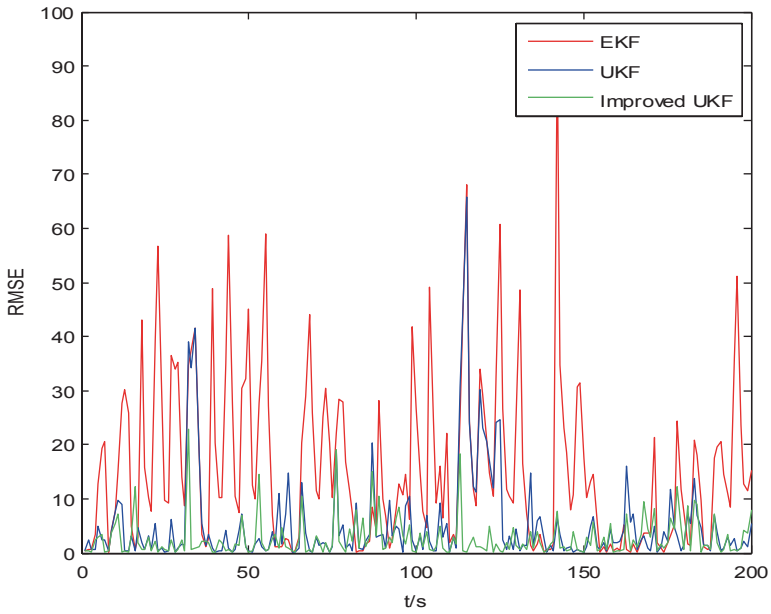


Fig. 65.5 Root mean square error (*RMSE*) of three algorithms. *EKF* extended Kalman filter, *UKF* unscented Kalman filter

Table 65.1 Computation times and mean value of RMSE

Algorithm	Computation times (s)	Mean value of RMSE
EKF	0.0324	9.795
UKF	0.0542	5.2376
Improved UKF	0.0063	2.8645

EKF extended Kalman filter, *RMSE* root mean square error, *UKF* unscented Kalman filter

65.5 Conclusion

In this chapter, an improved UKF algorithm is given which used the QR decomposition to update the squarer root of the covariance matrix P by Cholesky factor updating. Three nonlinear Kalman filter algorithms have been analyzed. From the simulation results, we can conclude that all of them can accomplish the estimation task, but the EKF has the worst accuracy, and the UKF needs a longer computation time, while the improved UKF achieves the best performance on the basis of the standard UKF.

References

1. Tilton AK, Ghiotto S. A comparative study of nonlinear filtering techniques. Proceeding of the 16th international conference on information fusion. Computer and Digital Engineering, IEEE. 2013;9–12.
2. Kalman RE, Bucy RS. New results in linear filtering and prediction theory. J Fluids Eng. 1961;83(1):95–108.
3. Julier SJ, Uhlmann JK. A new approach for filtering nonlinear system. Proceedings of the 1995 American Control Conference, IEEE. 1995;1628–1632.
4. Saadeddin K. Estimating vehicle state by GPS/IMU fusion with vehicle dynamics. J Intell Robot Syst. 2014;74 (1):147–172.
5. Cui W, Zhu K. Comparison of three filters in asteroid-based autonomous navigation. Res Astron Astrophys. 2014;14(3):329–344
6. Giannitrapani A, Scortecci F, Gatulli A. Comparison of EKF and UKF for spacecraft localization via angle measurements. Trans Aerosp Electron Syst. 2011;47(1):75–84.
7. Zhen L, Huajing F. Modified state prediction algorithm based on UKF. J Syst Eng Electron. 2013;24 (1):135–140.
8. Bistrovs V, Kluga A. The analysis of the UKF-based navigation algorithm during GPS Outage. Elektron Elektrotechnika. 2013;19(10):13–16.
9. Wang P, Xie W, Liu Z, Li P. Performance evaluation of several methods for tracking a ballistic object. J Shenzhen Univ Sci Eng. 2012;29(5):392–398. (In Chinese).
10. Liu C, Shui P, Li S. Unscented extended Kalman filter for target tracking. J Syst Eng Electron. 2011;22(2):188–192.
11. Cheng S. Unscented transformation and unscented Kalman filtering. Comput Eng Appl. 2008;44(24):25–35. (In Chinese).
12. Lin XL, Jiang YL. QR decomposition and algorithm for unitary symmetric matrix. Chin J Comput. 2005;28(5):817–822. (In Chinese).
13. Jing M. A comparison of three nonlinear filters for reentry ballistic target tracking. J Inf Comput Sci. 2012;9(17):5283–5290.

Chapter 66
Comparison of UAV Target Tracking Techniques

Dajian Li, Huixia Liu, Yangxia Dong, Qingbiao Xi and Ruofei He

Abstract Research on the unmanned aerial vehicle (UAV) target tracking system and the development of the UAV target tracking methods are discussed. A mean shift-auxiliary particle filter (M-APF) algorithm is proposed. It adopts APF as its main framework of the tracking algorithm; then, the mean shift is applied to calculate the offset of a few auxiliary particles and move them to the local optimum position of the observed values. Simulation results show that the calculation cost and robustness of M-APF are superior to mean shift algorithm and APF algorithm and satisfy the UAV target tracking requirements.

Keywords UAV · M-APF algorithm · Mean shift · Auxiliary particle filter · Target tracking

66.1 Introduction

As an indispensable military equipment in modern warfare, the unmanned aerial vehicle (UAV) system has been widely applied. Among the multiple tasks, the information reconnaissance and the war field surveillance are major combat tasks for the UAV system. An UAV can be used for strategic campaign or tactic reconnaissance [1]. It is capable of flying above the enemy target and scout both on day time and at night. It can transmit the real-time image and other information of the target to the command center so that the commander is able of making up combat plans based on the information; on the other hand, by the UAV system's mission evolving from only reconnaissance to reconnaissance and assaulting, tracking the

H. Liu (✉) · Y. Dong · Q. Xi
School of Automation, Northwestern Polytechnical University, 710072 Xi'an, China
e-mail: liuhuixia_nwpu@126.com

D. Li · H. Liu · Q. Xi · R. He
No. 365 Research Institute, Northwestern Polytechnical University, 710065 Xi'an, China
e-mail: liuhuixia_nwpu@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_66

enemy target steadily and accurately has become a primary precondition for UAV to achieve various tasks; therefore, a real-time tracking algorithm adaptable to complicated war field environment is of great importance for the UAV applications and development [1].

66.2 Tracking Algorithms

In the visual tracking system, the selection of the target tracking model would directly affect the accuracy of the tracking result. Generally, the color, the pattern, or the gradient of the target is often selected as a feature of the target for tracking. In different application scenarios, the feature selection varies. In this chapter, all the algorithms use the color of the target as the tracking feature. After the region of tracking is chosen, the histogram model of the color for this region is built as the target tracking model.

66.2.1 Mean Shift Algorithm

The mean shift algorithm is a nonparametric probability estimation algorithm [2, 7]. It applies to the gradient optimization method to the probability density function (p.d.f.) of the feature pixel. By recursively optimization, it converges to the local maximum of the p.d.f. rapidly by which the target can be tracked swiftly. The mean shift algorithm is able of tracking the target in real time and adapting to some levels of target morphing and rotating.

In order to apply the mean shift algorithm, the target model needs to be constructed first, that is, for the initial image, compute the feature value probability in the feature space of all the pixels in the target region. After that, the candidate models for the current image need to be built. Similar to the target model, it is to compute the feature value probability in the feature space of all the pixels in the candidate region for the current image. Finally, with the similarity measurement function, the similarity between the target model in the initial image and the candidate model in the current image is calculated. The average offset vector can be obtained by maximizing the similarity function. This average offset vector is the vector that the target is moving from the initial position to the current position. With the convergence attribute of the average offset vector, the computation of the average offset vector recursively will make the vector finally converge to the true position of the target in the current frame so that the target is successfully tracked. The mean shift algorithm is a data-driven tracking algorithm.

66.2.2 Auxiliary Particle Filter Algorithm

The particle filter algorithm [3, 5, 6, 9] is one of the most popular mathematical tools in solving the nonlinear problems in the target tracking area. It provides a probability model framework of target state prediction by the posterior p.d.f. Based on the importance sampling and the resampling technique, this algorithm is a model-driven tracking algorithm. The classic particle filter method needs a huge amount of dense samplings to achieve a high accuracy of target estimation; however, the algorithm often encounters the problem of lacking divergent particles. Based on the basic particle filter, the auxiliary particle filter (APF) algorithm includes the current measurements to the importance distribution to provide an importance distribution with relatively even weights of the particles. The divergence of the particles is improved, and the problem of lacking divergent particles is solved [3]. It is suitable for the target tracking with the nonlinear model and non-Gaussian noise.

The joint p.d.f. $p(x_k, i \mid z_{1:k})$ can be described as the following by the Bayes' rule:

$$\begin{aligned} p(x_k, i \mid z_{1:k}) &\propto p(z_k \mid x_k) p(x_k, i \mid z_{1:k-1}) \\ &= p(z_k \mid x_k) p(x_k \mid i, z_{1:k-1}) p(i \mid z_{1:k-1}). \\ &= p(z_k \mid x_k) p(x_k \mid x_{k-1}^i) \omega_{k-1}^i \end{aligned} \tag{66.1}$$

$p(x_k, i \mid z_{1:k})$ is the nonlinear model in UAV video target tracking; the particle is not able to be sampled directly, so the importance density function satisfies the following proportional relationship:

$$q(x_k, i \mid z_{1:k}) \propto p(z_k \mid \lambda_k^i) p(x_k \mid x_{k-1}^i) \omega_{k-1}^i, \tag{66.2}$$

where λ_k^i is the probability feature description of x_k given x_{k-1}^i. It can be the expectation $\lambda_k^i = E[x_k \mid x_{k-1}^i]$ or a sample $\lambda_k^i \sim p(x_k \mid x_{k-1}^i)$. In the experiment, we set the state transferring noise ω_k as a zero-mean Gaussian noise. Thus, the auxiliary particle collection can be chosen based on $\lambda_k^i \sim p(x_k \mid x_{k-1}^i)$. Let

$$\lambda_k^i = x_{k-1}^i + B \cdot no_k^i, \tag{66.3}$$

where x_{k-1}^i is the sampled particle based on the $(k-1)$th tracking result, no_k^i follows a Gaussian distribution with mean as 0 and variance as 1.

66.2.3 Mean Shift of a Regional Color Distribution Combined with the Particle Filter Algorithms

Mean shift of a regional color distribution combined with the particle filter algorithms [4] combines features of the mean shift algorithm and the particle filter algorithm. It designs a tracking algorithm with the meanshift framework. In order to

avoid the local minimum or maximum problem, the particle filtering algorithm is applied in the optimization phase to search for the true local optimum value. In this algorithm, only a few particles are used, and the computational cost of the particle filter algorithm is greatly reduced.

The workflow of the algorithm is as the following: The sampled particle at time k is composed of two types—one is selected from the tracking result with a fixed probability, and the other one is obtained by the weighted particle at time $k-1$w. Then, the states of the particles at time k are predicted, and the weights of the particles are calculated. Next, the mean shift algorithm drives the tracker and obtains the tracking result. The algorithm compares the Bhattacharyya distance between the particle prediction result and the tracking result of the target model. If the prediction is better than the tracking result, move the tracker to the particle prediction position and adjust the tracking region.

66.2.4 Mean Shift-Auxiliary Particle Filter Algorithm

Because of the complication of the UAV reconnaissance environment and the movement feature of the tracking target, the mean shift-auxiliary particle filter (M-APF) algorithm applies the APF algorithm as the framework of the tracking algorithm. It selects the tracking region from the initial image and builds the target model. The samples are sampled from the particle collection $\{x_k^j, i^j\}_{j=1}^N$ based on the importance density function $q(x_k, i \mid z_{1:k})$. The mean shift algorithm is then used as deterministic optimization to move each sample particle to the local maximum of the posterior density function so that the sample particles is approaching to the true location area of the target. In this way, the application of only fewer particles can achieve the same result as the conventional APF algorithm. The real-time performance of the algorithm is then improved.

66.3 Simulation Analysis

In this chapter, the video obtained from the UAV payload is used for the experiment. To verify the effect of the M-APF tracking algorithm proposed by this chapter, we employ four tracking algorithms for comparison: the mean shift algorithm, the APF algorithm, the mean shift of a regional color distribution combined with the particle filter algorithms, and the M-APF algorithm. We compare the real-time performances and the robustness of the four algorithms. The APF algorithm uses $N=200$ particles. The mean shift of a regional color distribution combined with the particle filter algorithms uses $N=20$ particles. The M-APF algorithm uses $N=10$ particles.

The hardware and software environments of this chapter are CPU: Q9500 2.83 GHz, memory: 4G, OS: Windows, development platform: Visual Studio 2008, and developing language: Open CV and C++.

66.3.1 Real-Time Performance

In this chapter, we use 300 frames of an image captured from the real UAV video to do the simulation. The resolution of the image is 614×467. The tracking window is selected manually in the initial frame. In about the 220th frame, some covering shadows occur in the tracking scenario, and the target begins to be covered by the shadow. In the 255th frame, the target is severely covered. In the 275th frame, the target leaves the shadow.

Table 66.1 provides the statistic results of the time cost for the four algorithms tracking targets. We divide the 300 frames of the image into six groups: 1st to 50th frame, 51st to 100th frame, 100th to 150th frame, 150th to 200th frame, 200th to 250th frame, and 251st to 300th frame; calculate the average time cost of each group and the whole video and list them in Table 66.1.

By the simulation result in Table 66.1, the M-APF algorithm proposed by this chapter spends less time than the APF algorithm and the mean shift of a regional color distribution combined with the particle filter algorithms for all the groups. For example, for the 151st to 200th frame, the time costs of the mean shift algorithm, APF algorithm, and mean shift of a regional color distribution combined with the particle filter algorithms are 1.69, 2.14, and 2.26 ms, respectively. Meanwhile, the algorithm proposed by this chapter spends 1.84 ms. For the whole tracking process, the first three algorithms spend 1.68, 2.13, and 2.29 ms per frame, while the M-APF algorithm spends merely 1.84 ms per frame. Comparing with the mean shift algorithm, the time cost of the M-APF algorithm is slightly higher. However, comparing with APF algorithm and the mean shift of a regional color distribution combined with the particle filter algorithms, the M-APF algorithm raises the processing speed significantly by using fewer particles. It has a better real-time performance than the APF algorithm and the mean shift of a regional color distribution combined with the particle filter algorithms. It can satisfy the real-time performance requirement of the UAV visual tracking system and is available to implement the real-time mobile object tracking for the UAV platform.

Table 66.1 Average time cost of the four algorithms

Frame	Mean shift (ms)	APF (ms) (N=200)	Ref [8] (ms) (N=20)	This chapter (ms) (N=10)
1–50	1.67745	2.13291	2.30254	1.84693
51–100	1.67996	2.13684	2.359	1.84167
101–150	1.6825	2.13285	2.39027	1.84674
151–200	1.68669	2.13533	2.25862	1.84688
201–250	1.67894	2.13218	2.21566	1.84393
251–300	1.67876	2.13297	2.20185	1.84114
Average time cost	1.68072	2.13385	2.28799	1.84455

APF auxiliary particle filter

66.3.2 Robustness

We show the simulation result of the 1st, 187th, 255th, and 275th frames. Figures 66.1, 66.2, 66.3, and 66.4 provide the tracking result of the mean shift algorithm, the APF algorithm, the mean shift of a regional color distribution combined with the particle filter algorithms, and the M-APF algorithm. For each group of the

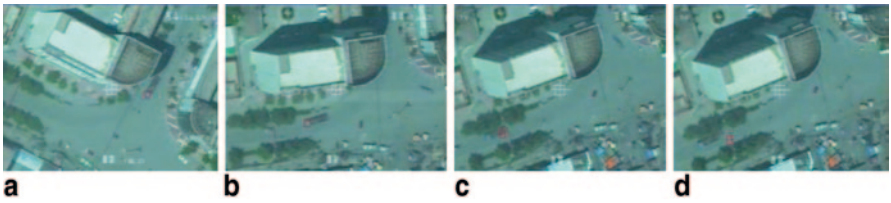


Fig. 66.1 Tracking result of the mean shift algorithm

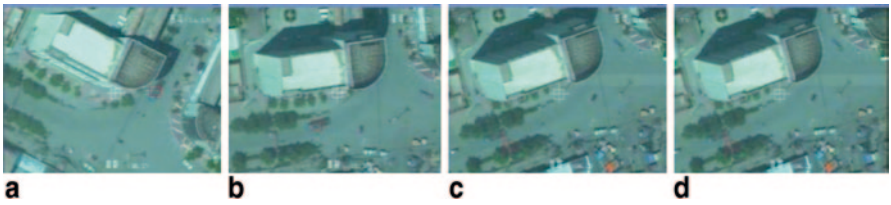


Fig. 66.2 Tracking result of the APF algorithm

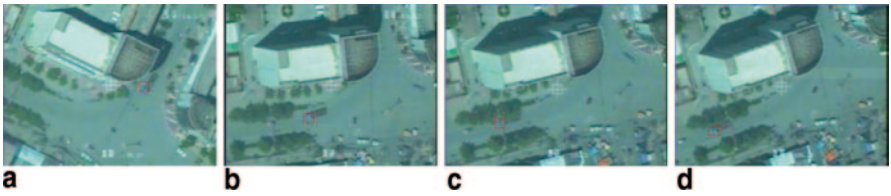


Fig. 66.3 Tracking result of the mean shift of a regional color distribution combined with the particle filter algorithms

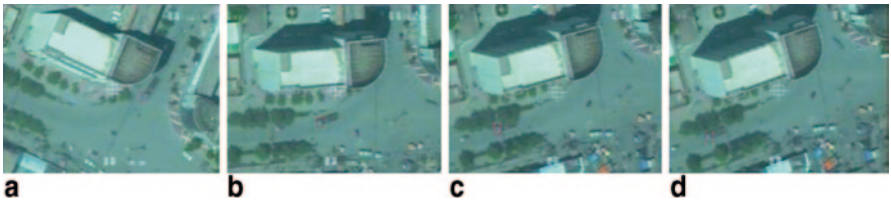


Fig. 66.4 Tracking result of the M-APF algorithm

figure sequence, figure part a is the initial frame. The region marked by a red rectangle is the selected tracking area. Figure parts b, c, and d are the tracking results of the 187th, 255th, and 275th frames of the four algorithms where the tracked region is marked with the red rectangle.

Comparing figure parts a and b in each figure group, the rectangles in all the figures can lock the tracking target, which implies that when the tracking target is experiencing some level of rotating and small background illumination changes, all the algorithms can track the target stably. By figure part c of all the groups, when the video moves to the 255th frame and the target is being covered, the mean shift algorithm and the APF algorithm happen to lose the target slightly. By figure part d, the mean shift algorithm and the APF algorithm lose the target completely when the target leaves the covered area. However, in Figs. 66.3d and 66.4d, the target is still tracked stably, which means that the mean shift of a regional color distribution combined with the particle filter algorithms and the M-APF algorithm are able of keeping stable track even when the target is covered.

On the basis of analysis of the real-time performance and the robustness of the four algorithms, the M-APF algorithm proposed by this chapter is slightly weaker than the mean shift algorithm in the real-time performance but is better in the robustness aspect. Comparing the APF algorithm with the mean shift of a regional color distribution combined with the particle filter algorithms, the M-APF algorithm is significantly better in real-time performance; therefore, the algorithm proposed by this chapter is superior to the rest three algorithms by the general performance. Specifically, the algorithm proposed by this chapter is capable to track the target stably and accurately when the target is covered and it has a good real-time performance. It can meet the requirement of real-time performance and robustness in the practical application of an UAV target tracking system.

66.4 Conclusion

With the complicated ground environment and high mobility of the target, the UAV mobile target tracking system often encounters the problem of poor real-time performance and the problem of losing a target when it is covered. This chapter proposes a real-time UAV target tracking algorithm based on the M-APF algorithm and compares it with the mean shift algorithm, APF algorithm, and mean shift of a regional color distribution combined with the particle filter algorithms. The result shows that the M-APF algorithm provides a good solution to solve these problems. It can perform in real time and track the target when it is experiencing some level of rotating, scaling, and small background illumination changes. It can also track the target stably and accurately when the target is covered by shadows. It may meet the requirements of real-time performance and robustness for the UAV platform target tracking system; however, the tracking target selected in this chapter is of small size and high mobility. In the future, we will focus on the rapidly rotating targets with a larger size.

References

1. Unmanned Aerial Vehicle (UAV) roadmap 2005–2030. Office of the Secretary of Defense. USA 2005.
2. Zhang L. Research of video object tracking method. Hefei: University of Science and Technology of China; 2009.
3. Hu S, Jing Z. Particle filter theory and its application. Beijing: Science; 2010 (In Chinese).
4. Deguchi K, Kawanaka O, Okatani T. Object tracking by the mean-shift of regional color distribution combined with the particle-filter algorithms. Proceedings of the 17th International Conference on Pattern Recognition. Cambridge, UK: IEEE Computer Society; 2004. p. 506–9.
5. Wang X, Fang L, Chong Z. Research on video vehicle tracking algorithm based on Kalman and particle filter. J Image Graph. 2010;15(11):1615–22.
6. Isard M, Blake A. Condensation-conditional density propagation for visual tracking. Int J Comput Vis. 1998;29(1):5–28.
7. Tada K, Takemura H, Mizoguchi H. Robust tracking method by Meanshift using Spatiograms. SICE Annual Conference 2010, 18–21 Aug:1985–1988. 2010.
8. Chen Z-X, Zhao X-J, Nie S-D. Research and applications of mean shift method in image processing. Chin J Med Phys. 2010;27(6):2244–9.
9. Sullivan J, Rittscher JR. Guiding random particles by deterministic search. Proceedings of ICCV'01. [S. I.]: IEEE Press; 2001. p. 323–30.

Chapter 67
Fuzzy Algorithm-Based Manufacturing Equipment Classification

Yujie Shi and Dehui Zhai

Abstract In order to improve the efficiency of the algorithm for fast classification and retrieval of the manufacturing equipment, a concept of manufacturing feature vector (MFV)-based manufacturing equipment description is introduced. First, the type, size and precision are defined as three attributes of the MFV; and second, an extended fuzzy C-means clustering algorithm has been amended to fulfill the aim of manufacturing equipment classifying. The algorithm can efficiently and accurately calculate the range of the optimal grouping number; finally, a sample of 50 manufacturing equipment classifications is used to verify the effectiveness and accuracy of the algorithm.

Keywords Fuzzy C-means clustering · Manufacturing feature vector · Equipment grouping

67.1 Introduction

In the manufacturability evaluation, process planning and production planning, suitable manufacturing equipment needs to be selected by the machining feature. In the networked manufacturing environment, the manufacturing resources have the characteristic of having many kinds of varieties, and a large number and a wide range of processing capacity, which is time-consuming for selecting the suitable equipment for each machining feature. When complex parts contain more machining features, it makes the problem even more difficult [1]; therefore, we must research efficient and accurate methods of equipment grouping based on the manufacturing features (MFs).

The optimal grouping of manufacturing equipment is actually a pattern recognition and clustering problem [2]. Currently, the manufacturing equipment grouping methods are mainly based on fuzzy clustering methods. Among numerous fuzzy

Y. Shi (✉) · D. Zhai
School of Mechanical Engineering, Beijing Institute of Technology, 100081 Beijing, China
e-mail: shiyujie92@126.com

D. Zhai
e-mail: zdh20081461@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_67

clustering methods, fuzzy C-means clustering (FCM) algorithm [3] can obtain the degree of uncertainty of samples belonging to various groups fit for manufacturing equipment grouping; however, FCM algorithm requires setting the number of groups. It cannot solve the optimization of the grouping number. To overcome the limitations of the FCM algorithm, Shi Xudong et al. [4] proposed the expanded fuzzy C-means clustering method (EFCM); but this method uses the traversal method in the optimization of grouping number, which results in the complexity of algorithm.

As to the disadvantages of the above grouping methods, EFCM clustering algorithm and the nonergodic group optimization method as well as the grouping number optimization method are combined as manufacturing equipment grouping algorithm. At last, an example was verified.

67.2 Description of the Manufacturing Equipment Feature Vector

Description of the manufacturing equipment is the base of manufacturing equipment grouping, which provides the basic support for equipment evaluation, retrieval, management, and portfolio. As a minimum processing unit contains a wealth of manufacturing information [5], MFs are the foundation of describing manufacturing equipment. MFs are the basic units of describing geometry shape processing activities, which are combined by the geometry shape and the process constraints [6].

Because of different size and machining accuracy requirements of the MFs, the same type of geometric features may require different manufacturing equipment; therefore, the manufacturing equipment grouping based on MFs has to consider not only GF but also PF (Precision Feature) and DF (Dimension Feature). Thus, the manufacturing equipment vector can be described mathematically as follows:

$$e_i = \left(e_{i,1}, e_{i,2}, \cdots\cdots, e_{i,s}, b_i, t_i\right) \quad i = 1, 2, \cdots\cdots, n \tag{67.1}$$

where $e_{i,k}$ represents the feature capability index of the i-th equipment processing the k-th MF, b_i represents precision capability index of the i-th equipment, t_i represents dimension capability index of the i-th equipment, and n means the total equipment number.

The corresponding relationships between MFs symbolic representation and manufacturing equipment vector are described in Table 67.1.

Table 67.1 Correspondence table of manufacturing features (MFs) and equipment feature matrix

MF	Plane	Stepped surface	Curved surface	Cylinder	Slot	Ring groove	Slot series	Hole	Hole series	Thread
Symbolic representation	P	P*st*	C	C*e*	S	S*c*	S*t*	H	H*s*	H*t*
Feature vector	$e_{i,1}$	$e_{i,2}$	$e_{i,3}$	$e_{i,4}$	$e_{i,5}$	$e_{i,6}$	$e_{i,7}$	$e_{i,8}$	$e_{i,9}$	$e_{i,10}$

67.3 Fuzzy Grouping Algorithm of Manufacturing Equipment

FCM clustering algorithm is a fuzzy grouping method for samples, and it is often used for manufacturing equipment grouping. EFCM clustering algorithm based on the FCM algorithm will consider both the tightness within the group and the deviation among groups.

As to the defects of EFCM algorithm, this chapter uses an equipment grouping algorithm based on the MFV and initial grouping center selecting strategy to calculate the range of optimal group number. This algorithm can find suitable manufacturing equipment fast and efficiently.

The algorithm steps are given below:

Step 1: The initialization step, and initial grouping number $m=1$.
Step 2: The initial grouping center selecting strategy.

If the grouping number $m=1$, randomly select a device as the initial group center G_1. If the grouping number $m=2$, pick two Euclidean distances in the Euclidean vector space device characterization as the first two cluster centers G_1 and G_2. If the grouping number $m>2$, we can set a collection before the m-1 initial group center as $\{G_1,\ldots,G_{m-1}\}$; characterization of each device in order to calculate the vector (where n represents the total number of devices) and the distance between $D_{i,s}$, where

$$\left\{D_{i,s}=\left\|M_i-G_{i,k}\right\|\middle| s=1,\ldots,m-1\right\}. \tag{67.2}$$

Choose the longest distance from the minimum value of the distance between all of the devices and the grouping centers of the front of (m-1) devices, and the describing feature vector of the corresponding devices is taken as the grouping center of the m^{th} group G_m.

That is, if

$$\min(D_{i,1},\ldots,D_{i,m-1})=\max\left\{\min\left\{D_{1,1},\ldots,D_{1,m-1}\right\},\ldots,\min\left\{D_{n,1},\ldots,D_{n,m-1}\right\}\right\},$$

then $G_m=M_i$.

Step 3: Calculate the number of grouping m which corresponds to the packet comprehensive evaluation results and the clustering R according to the EFCM method and the choosing of the initial grouping center (see step 2).

1. Calculate the membership of each device to each packet as U_{ij}, and U_{ij}expresses the ratio of the geometric distance between the manufacturing device M_i and the grouping center G_j and the geometric distance between manufacturing device M_i and all of the grouping centers. Select the device M_i which is included in the group that has the biggest membership, and we can set the equipment into this group, then we will complete a grouping.

$$u_{ij} = \frac{\left(\frac{1}{e^2(m_i, G_i)}\right)^{1/(f-1)}}{\sum_{j=1}^{m}\left(\frac{1}{e^2(M_i, G_j)}\right)^{1/(f-1)}}, \tag{67.3}$$

where $e^2(M_i, G_j)$ indicates the distance between the device characterization *Mi* and the description vector of the grouping center G_j in Euclidean space; *f* represents the fuzzy index, which should be greater than 1.

2. According to the membership calculated from step 3 (1), we can recalculate the grouping center G_j of each group.

$$G_j = \frac{\sum_{i=1}^{n_j}(u_{ij})^f M_i}{\sum_{i=1}^{n_j}(u_{ij})^f}. \tag{67.4}$$

3. Calculate the processing characteristics of the membership of each group. The processing characteristics k of the membership g_{kj} in each group *j* represents a group describing manufacturing apparatus for weighted feature vectors and a weighting of all groups of the machining apparatus characterized by getting the ratio of feature vectors which is described.

$$g_{kj} = \frac{r_{kj}}{\sum_{j=1}^{m} r_{kj}}, \tag{67.5}$$

where R_{kj} expresses the *j* group which has the characteristics *k* of the device manufacturing apparatus having the characteristics described in the processing and weighting vectors.

4. Calculate the comprehensive evaluation index clustering *R*, and *R* describes the reciprocal duplication of the average deviation of equipment λ and processing characteristics of the repeated index *r*.

$$R = \lambda + \frac{1}{r}. \tag{67.6}$$

5. Repeat step 3 from (3) to (6) until the clustering comprehensive evaluation index R does not decrease, then get the clustering comprehensive evaluation index which corresponds to *m*.

Step 4: Determine the optimal number of groups setting range.

The relationship between the clustering comprehensive evaluation function R and the number of groups m is shown in Eq. (67.7).

$$R(m) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j + \frac{1}{\frac{1}{m}\sum_{j=1}^{m}\frac{1}{n_j}\sum_{i=1}^{k}(g_{ij})^2}. \tag{67.7}$$

In order to research $R(m)$ which is changing with m and confirm the range of the optimal number of setting groups, we can use qualitative, quantitative and comprehensive analysis of qualitative and quantitative Eq. (67.7).

$R(m)$ can be expressed by function $f_1(m) = h/m + m^t$ approximately, where $t \in [1,2]$, $h \in \left[0, \lambda_1^1\right]$;

$m = (h/m)^{\frac{1}{t+1}}$ is the only one solution of the function $f_1(m) = h/m + m^t$ in the range of $\left[1, +\infty\right)$, and this is the optimal number of setting groups.

Step 5: Update the grouping number m.

Update m=m+1, if $m \in \left[1, \lambda_1^1\right]$, repeat steps 2 and 3, and determine the number of packets m corresponding to the grouping result and comprehensive evaluation index R; if $m \notin \left[1, \lambda_1^1\right]$, then we should choose the smallest cluster comprehensive evaluation index corresponding packet as the optimal grouping. And optimal grouping results consist of three parts: the optimal number of packet groups, manufacturing equipment which is included in each group and for each machining feature which belongs to membership of each group.

67.4 Examples Verification

The grouping algorithm can be validated by the following example. A workshop has 50 machine tools. The feature vector can be used to describe the ability of the machine tools. Table 67.2 gives some description of the MFV.

67.4.1 Analysis of Results

This chapter calculated all possible clustering evaluation index R corresponding to the grouping number of 50 manufacturing equipments. With the changes in grouping numbers, clustering evaluation index R was shown in Fig. 67.1.

The value of evaluation index R was {38,17.5,14,12.5,14.5,16,16.5,17,18} when the grouping numbers are among {1,2,3,4,5,6,7,8,9}.

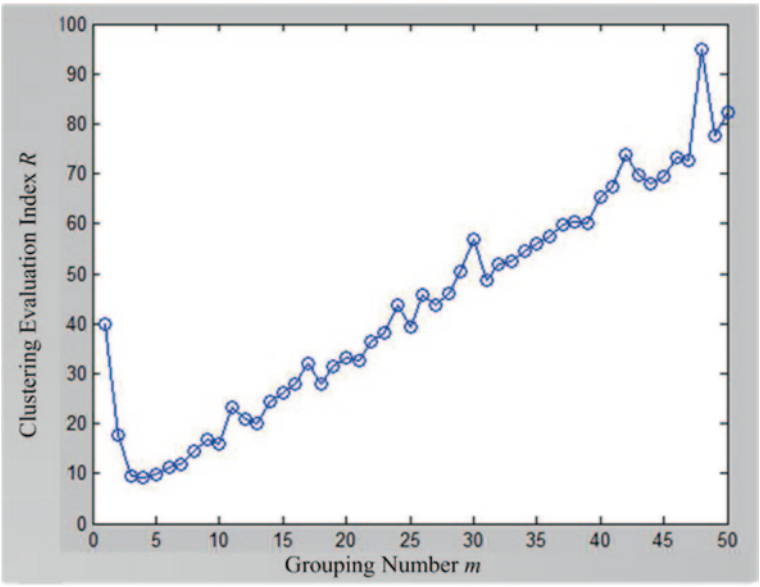


Fig. 67.1 Relationship of grouping number and grouping effect

As the largest similarity in the group and the biggest difference between two groups are the best, the clustering evaluation index R is smaller and better. When the grouping number is four, corresponding clustering evaluation index R was optimal.

67.4.2 Contrast Algorithm

The article uses MATLAB 7.0 in Intel Core i3-2310M platform to achieve the algorithm above. An example of Table 67.2 is calculated by the standard FCM algorithm, EFCM algorithm, and manufacturing equipment grouping algorithm based on MFs. The efficiency and the comparative results are shown in Table 67.3.

67.5 Conclusion

This chapter has presented the manufacturing equipment grouping algorithm based on MFV. The algorithm uses the FCM technology to device the similar processing units together into a class according to description of manufacturing equipment feature vector. Due to the increasing initial grouping center selecting strategy and the non-ergodic group optimization techniques, it may improve the efficiency of grouping and retrieving manufacturing equipment. FCM algorithm, EFCM algorithm and the algorithm of this chapter have been adopted to group the 50 manufacturing

Table 67.2 List of MFV

Equipment no.	Equipment name	MF											
		P	Pst	C	Ce	S	Sc	St	H	Hs	Ht	B	t
1	Lathe	0	1	0	1	0	1	0	0.5	0	1	0.5	0.5
2	Milling machine	1	0	0	0	1	0	0	0	0	0	0	1
3	Lathe	0	1	0	1	0	1	0	0.5	0	1	1	1
4	Lathe	0	1	0	1	0	1	0	1	1	1	1	1
5	Cylindrical grinder	0	0	0	1	0	1	0	0	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
46	Milling machine	1	0	1	0	1	0	0.5	0	0	0	0.5	1
47	Drilling machine	0	0	0	0	0	0	0	1	0	1	0	0
48	Boring machine	0	0	0	0	0	0	0	1	0	0	1	1
49	Special machine	0	0	0	0	0	0	1	0	0	0	1	1
50	Milling machine	1	0	0	0	1	0	0	0	0	0	0.5	0

MF manufacturing feature

Table 67.3 Comparison table of different grouping algorithms

Algorithm	Example	Search range	Algorithm time (s)	Clustering evaluation index *R*
FCM algorithm	50 manufacturing equipments	1–50	15.6	15.5901
EFCM algorithm		1–50	52.5	12.7982
Algorithm of this chapter		1–9	1.7	12.7982

FCM fuzzy C-means clustering, *EFCM* expanded fuzzy C-means clustering method

equipments of the example, and the result shows that the efficiency of grouping through the algorithm of this chapter is significantly higher than the other two.

References

1. He W. Research on technique of the resource optimization deploying for networked manufacturing. Xi'an: Northwestern Polytechnical University; 2007 (In Chinese).
2. Patrick CHM, Keith CCC, Xin Y. An evolutionary clustering algorithm for gene expression data analysis. IEEE Trans Evolut Comput. 2006;3(6):112–8.

3. Xiao M, Wang X, Wang L. Web resource recommendation method based on intuitive fuzzy clustering. J Comput Appl. 2012;9(10):2480–2.
4. Shi X, Li F, Li J. Virtual manufacturing modeling for manufacturability evaluation. China Mech Eng. 2002;13(17):1483–5 (In Chinese).
5. Ge C, Qiao L. Manufacturing feature information modeling and instantiation. Comput Integr Manuf Syst. 2010;16(12):2570–6 (In Chinese).
6. Ge C, Qiao L. Machining feature recognition based on geometry and tolerance information. Comput Integr Manuf Syst. 2010;2(10):256–62 (In Chinese).

Part VI
Signal Detection and Image Processing

Chapter 68
Vibration Test of a Dynamic Target Based on Stereo Machine Vision

Yingjie Dai, Wei Wang, Shijie Deng and Xujun Su

Abstract The position and posture measurement problem for a dynamic target is widely studied by researching the testing and calibration model of the vertical optical axis binocular vision, deducing the lineup structure of the cameras and the positional relationship between the cameras and the target, then establishing the noncontact vibration test system of a dynamic target. The test system uses double area array charge coupled devices (CCDs) to capture vibration images of the target. The system extracts the pixel position information to calculate the spatial coordinates of the target point; then, according to the target pixel position changes in the whole process, it calculates the displacement variation in three dimensions. Experimental results show that the measurement accuracy is about 0.1 mm, the relative error is less than 5 %, meeting the engineering test requirements.

Keywords Binocular vision · Vertical optical axis · Dynamic target · Vibration test

68.1 Introduction

As one of the most important fields in the modern photoelectric detection technology, the machine vision measurement technology has been widely used in the fields of industrial inspection, virtual reality, and the range test [1]. Machine vision measurement techniques include monocular vision technology [2], stereo vision measurement technology [3], and structured light vision measurement technology [4]. Monocular vision technology gets the motion information of the target through pictures from a single camera. It must be known that the information of the camera or objects is to restore the motion information of the target. Structured light vision measurement technology takes the projection point of the light source through a single camera. Projection angle and position of the light source are known.

Binocular stereo vision uses the disparity in the two images to restore the three-dimensional space coordinates [5]. In accordance with the arrangement of the

Y. Dai (✉) · W. Wang · S. Deng · X. Su
Mechanical Engineering College, 050003 Shijiazhuang, China
e-mail: dyj105.czzx@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_68

light-sensitive cells, the charge coupled device (CCD) can be divided into linear CCD and area array CCD. The advantage of the area array CCD is that it can receive a complete two-dimensional image, measure image intuitively, be highly efficient, and is not affected by the scan movement and position feedback link.

68.2 System Components and Measuring Principle

68.2.1 System Components

The measuring system consists of collecting and measuring parts. As shown in Fig. 68.1, the collecting part is for recording the target images. It consists of the computer, CORE, and CCD. The measuring part is to achieve the goal of measuring the displacement vibration. It consists of the APSM25A model multi-DOF displacement platform, the calibration target, and the solid circle array target.

68.2.2 Measuring Principle

Without considering the camera distortion, the linear camera model is the pinhole model. According to the principle of the pinhole model, the size of the measured object is *Y,* image height is *y,* object distance is *L,* image distance is *l,* and focal distance is *f*.

From the lens imaging formula, we can get $1/L + 1/l = 1/f$.

Optical imaging relationship is $y=\beta Y$,

where β is the magnification of the optical system $\beta = l/L = f/(L\text{-}f)$.

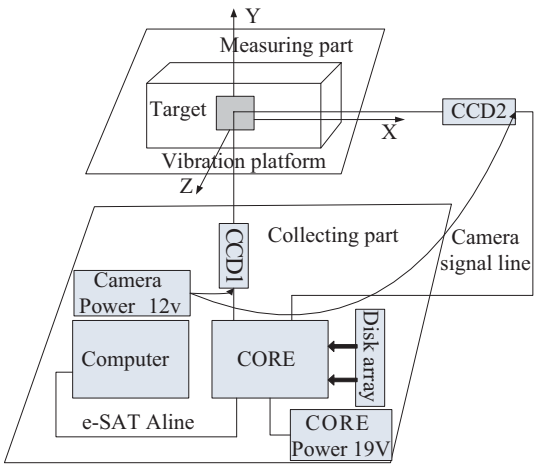


Fig. 68.1 Measuring system components

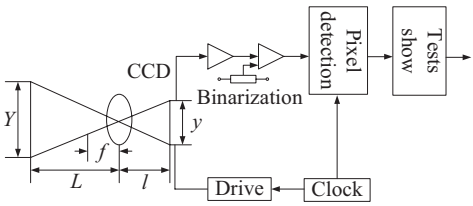


Fig. 68.2 Schematic representation of the linear CCD measuring system

Adjacent pixel spacing is d. Image height occupies the number of pixels which is N. So $y = Nd$. Through simultaneous equations, we can get: $Y = (L/f - 1)\,Nd$.

The linear CCD measuring system is shown schematically in Fig. 68.2.

68.3 Vertical Axis Binocular Stereo Vision Target Spatial Displacement Solver

68.3.1 Vertical Axis Binocular Stereo Vision Model

There are four coordinate systems in the binocular stereo vision model: (1) world coordinate system (X_w, Y_w, Z_w), as reference to describe the camera position of all the objects and environments; (2) camera coordinate system (X_c, Y_c, Z_c), with the camera optical center as origin O; (3) pixel coordinate system (u, v); and (4) image coordinate system (x, y).

As shown in Fig. 68.3a and b, point P is any point in space, point P′ is the projection of point P. In A image point is p_1, in $O_a - X_a Y_a Z_a$ the coordinate system is (X_1, Y_1, Z_1), in $o_1 - x_1 y_1$ the coordinate system is (x_1, y_1). In B image point is p_2, in

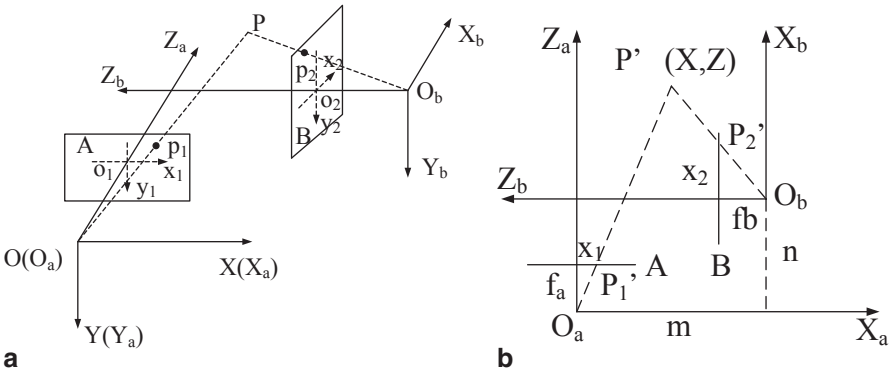


Fig. 68.3 **a** Calculation model of vertical optical axis. **b** x-z plane binocular stereo vision projection mapping

$O_b - X_b Y_b Z_b$ the coordinate system is (X_2, Y_2, Z_2), and in $o_2 - x_2 y_2$ the coordinate system is (x_2, y_2). So there are expressions as below:

$$\begin{cases} X_1 = \dfrac{x_1}{f} Z_1 \\ Y_1 = \dfrac{y_1}{f} Z_1 \end{cases} \quad \begin{cases} X_2 = \dfrac{x_2}{f} Z_2 \\ Y_2 = \dfrac{y_2}{f} Z_2 \end{cases} \quad \begin{cases} X = X_1 \\ Z_1 = n + X_2 \\ Z_2 = m - X_1 \end{cases} . \tag{68.1}$$

Through simultaneous three equations, we can get the world coordinates of point P.

$$\begin{cases} X = \dfrac{(nf + mx_2)x_1}{f^2 + x_1 x_2} \\ Y = \dfrac{y_1}{f}\left(n + \dfrac{(mf - nx_1)x_2}{f^2 + x_1 x_2}\right) \\ Z = n + \dfrac{(mf - nx_1)x_2}{f^2 + x_1 x_2} \end{cases} \tag{68.2}$$

68.3.2 Camera Space Lineup Structure

In order to ensure that the target movement is always in the field of two cameras' common view, we should place the camera to longer distance in principle. But with increasing distance from the camera to the target, the accuracy will reduce. Therefore, the lineup structure of the camera and its relationship with the target do the following derivation.

As shown in Fig. 68.4, we establish the *xoz* coordinate system. a_1, a_2, a_3, and a_4 make up the range of $p \times q$. The distance between camera a and the target is h. The distance between camera b and the target is l. Two cameras' field angles are both 2α.

In *xoz* coordinate system,

Linear equation of oa_1 is $y = \tan(\pi / 2 - \alpha)x$.

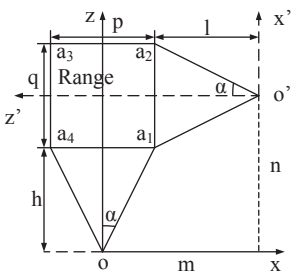


Fig. 68.4 Camera lineup structure diagram

Linear equation of oa_4 is $y=\tan(\pi/2+\alpha)x$.
Linear equation of $o'a_1$ is $y-n=\tan\alpha(x-m)$.
Linear equation of $o'a_2$ is $y-n=\tan(\pi-\alpha)(x-m)$.
Simultaneous four linear equations are as follows:

$$\begin{cases} h = y_{a1} = (n\cot\alpha - m)/(\cot\alpha - \tan\alpha) \\ l = m - x_{a1} = (m\cot\alpha - n)/(\cot\alpha - \tan\alpha) \\ p = x_{a1} - x_{a4} = 2(n - m\tan\alpha)/(\cot\alpha - \tan\alpha) \\ q = y_{a2} - y_1 = 2(m - n\tan\alpha)/(\cot\alpha - \tan\alpha) \end{cases}. \tag{68.3}$$

The achieved conditions are as follows:

$$\begin{cases} x_{a1} - x_{a4} \geq p \\ y_{a2} - y_{a1} \geq q \end{cases}. \tag{68.4}$$

On the basis of ensuring that the target movement is always in the field of two cameras' common view, we can achieve the highest possible accuracy.

68.4 Experiments and Data Analysis

68.4.1 Experimental Verification

In order to verify the effectiveness of the method, we do the vibration displacement experiments in the APSM25A multi-freedom platform. Experimental steps are as follows:

1. Prepare hardware, connect the computer, CORE, and place the camera in the right way. Fix the calibration target to the displacement test bench.
2. Turn on the computer, run the recording software CoreView. And adjust the camera focal length and the axis vertical contour.
3. Cameras take images with 6 to ~8 different poses of calibration target. Set the dynamic gray threshold of the original image, binarization processing, use the Canny operator to extract the circular outer contour [6], then use the Gaussian fitting to calculate the pixel coordinates of the circle center [7].
4. Calibrate cameras to get the internal and external parameters.
5. Record the reference image of the target's initial position. Run the displacement test bench, then record the work image of the target's motion state. Close the displacement test bench.
6. Do the image processing as step (3) to get the pixel coordinates of the circle center. Then, calculate the displacement of the target.
7. Change the vibration parameters. Make the next experiment. When done, exit the test system, finishing the equipment (Figs. 68.5 and 68.6).

Fig. 68.5 Experimental global overview

68.4.2 Experimental Results and Analysis

The measurement results are shown in Table 68.1. Each group setting is as described. Program solver results and relative errors are described in the table. Reference data are measured by the micrometer. Measurement results are obtained by

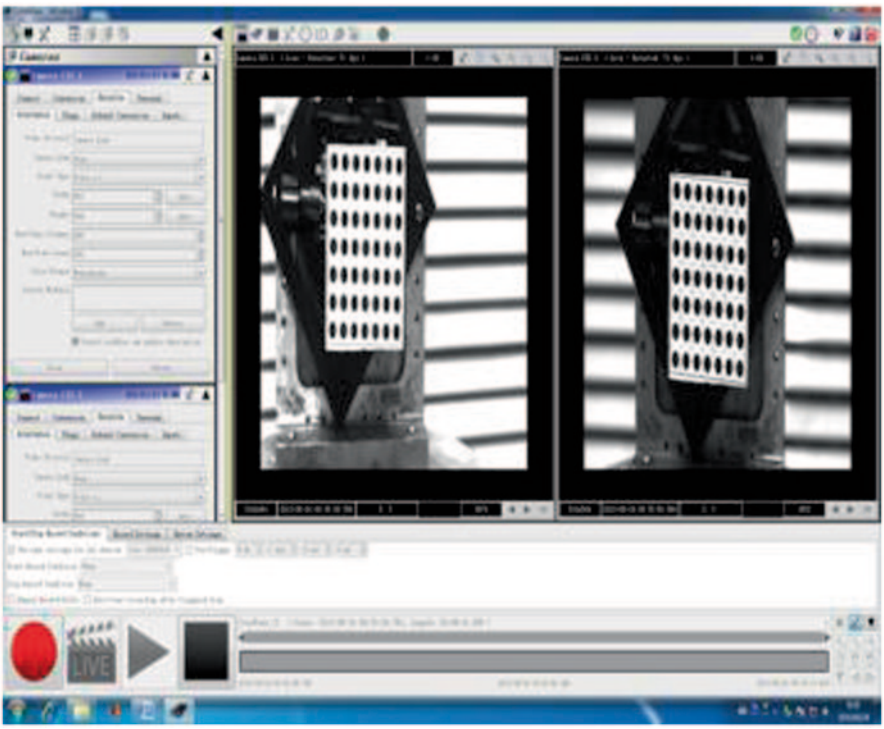


Fig. 68.6 CoreView software running calibration target images

Table 68.1 *XYZ*-axis vibration test results (Unit: mm)

Reference data	Measurement results			Measurement error			Maximum absolute error	Maximum relative error
X(Y, Z) displacement	X results	Y results	Z results	X error	Y error	Z error		
1.0000	1.0072	1.0068	9.9976	0.0072	0.0068	−0.0024	0.0072	0.72%
2.0000	2.0195	2.0137	1.9857	0.0195	0.0137	0.0143	0.0195	0.98%
3.0000	3.0305	3.0222	2.9812	0.0305	0.0222	−0.0188	0.0305	1.02%
4.0000	4.0304	4.0234	3.9749	0.0304	0.0234	−0.0251	0.0304	0.76%
5.0000	5.0467	5.0418	4.9523	0.0467	0.0418	−0.0477	−0.0477	0.95%
6.0000	6.0600	6.0528	5.9387	0.0600	0.0528	−0.0613	−0.0613	1.02%
7.0000	7.0775	7.0714	6.9296	0.0775	0.0714	−0.0704	0.0775	1.11%
8.0000	8.0812	8.0771	7.9147	0.0812	0.0771	−0.0853	−0.0853	1.07%
9.0000	9.0913	9.0868	8.9045	0.0913	0.0868	−0.0955	−0.0955	1.06%
10.0000	9.9072	9.9065	9.8997	−0.0928	−0.0935	−0.1003	−0.1003	1.00%

this algorithm. The error is between the reference data and the measurement results. Measured results greater than the reference data are positive, otherwise negative.

As can be seen from the data in Table 68.1 and Fig. 68.7, the maximum measurement error in the *X*-direction is −0.0928 mm, the maximum measurement error in the *Y*-direction is −0.0935 mm, the maximum measurement error in the *Z*-direction is −0.1003 mm. The average error in the *X*-direction is 0.0352 mm, and the variance is 0.0028 mm. The average error in the *Y*-direction is 0.0303 mm, and the variance is 0.0027 mm. The average error in the *Z*-direction is −0.0493 mm, and the variance is 0.0016 mm. Measurement accuracy is about 0.1 mm.

From the test results, it can be seen that as the displacement volume increases in the *XYZ*-axis direction, the absolute error increases, but the relative error decreases. The maximum absolute error in the *XYZ*-axis direction is less than −0.1003 mm, and the maximum relative error is less than 1.11%.

68.5 Conclusion

We use the stereo vision measurement technology with two area array CCD to establish the testing and calibration model of the vertical optical axis binocular vision. Deriving the lineup structure of the camera and its relationship with the target. On the basis of ensuring that the target movement is always in the field of two cameras' common view, we can achieve the highest possible accuracy. Given the limited space, omit the two cameras' calibration process. However, the minor deviation of the main point coordinate and focal length will cause considerable displacement calculation errors. Experimental results show that the measurement accuracy is about 0.1 mm, the relative error is less than 5%, meeting the engineering test requirements. The measurement system has the following features: compact, high

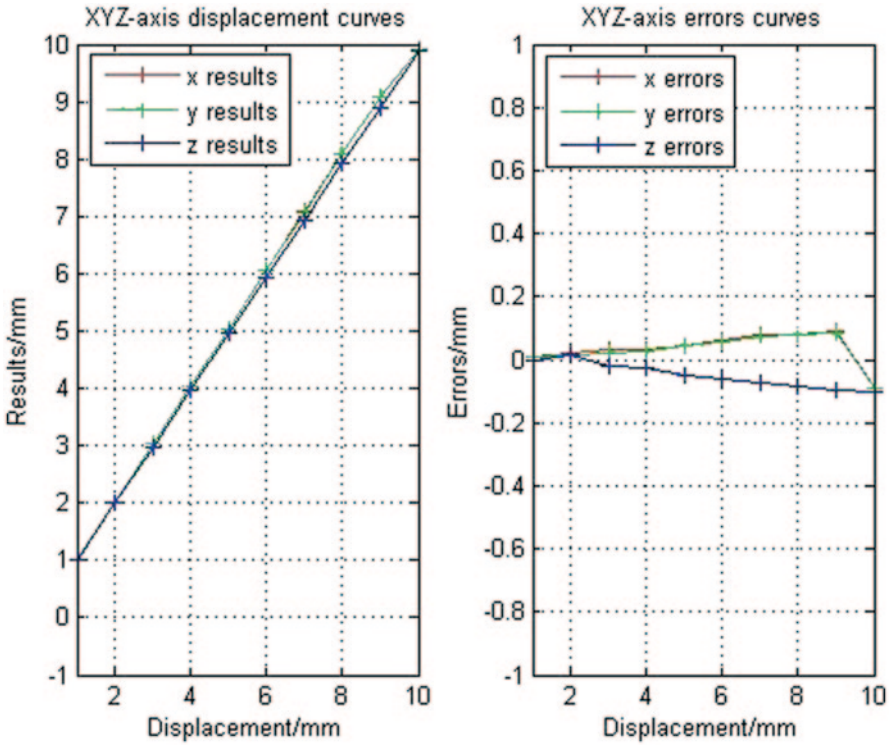


Fig. 68.7 *XYZ*-axis vibration test result curves

accuracy and applicability, and achieving the purposes of intellectualized test and data analysis process automation. It can be widely used to measure the objectives' three-dimensional space motion parameters and space gesture in engineering test areas, etc. By the way, effects of lens distortion and image quality factors affect the small displacement measurement accuracy obviously. Therefore, improving the measurement accuracy and the robustness of the system is the focus in the future research.

References

1. Xu Y, Jiang Y, Chen F. Stereo visual localization algorithm based on 2D-3D binocular motion estimation. Opto-Electron Eng. 2011;38(9):104–10.
2. Han Y, Zhang Z, Dai M. Monocular vision system for distance measurement based on feature points. Opt Precis Eng. 2011;19(5):1110–6.
3. An K, Ma J. A fast CCD measurement method used in rendezvous and docking. Opto-Electron Eng. 2011;38(6):1–6.
4. Xu L, Zeng D, Zhang Z. Error analysis in structured light system. Opto-Electron Eng. 2008;35(10):26–31.

5. Zou W-J, Zhu G, Wu X. Angle measuring system of laser based on array CCD. Opto-Electron Eng. 2006;33(10):91–5.
6. Wei H, Zhao B, Tang L, He P. Improvement of Canny algorithm and implementation in hardware. Optical Technique. 2006;32(2):263–6.
7. Liu G, Liu B, Chen F, Hu T. Study on the method of the accuracy evaluation of sub-pixel location operators. Acta Optica Sinica. 2009;29(12):3446–51.

Chapter 69
Fast Image Fog Removal Based on Gray Image Guided Filtering

Zhenyu Wang, Hang Li, Jing Teng and Xiaobo He

Abstract Under the condition of foggy weather, the videos and images obtained by the imaging devices which capture visible light will be severely degraded due to the low visibility of the scene, such as contrast reduction and color attenuation. The fog removal algorithm based on the dark channel prior now has yielded a great effect, but the algorithm has the characteristics of high time complexity and space complexity; thus, it does not have practicality. On the basis of the dark channel prior, we propose a fast method of image fog removal based on gray image guided filtering. Firstly, we estimate the atmospheric scattering model transmission through the dark channel information, and then adopt the gray image of the input mistily image to guide the transmission filtering to enable the optimization of rough transmission, namely, to maintain edges and smoothing regions; on this basis, the recovery image without fog can be obtained. Experiments demonstrate that the proposed algorithm in this chapter can effectively remove fog from a foggy image and increase the efficiency as a result.

Keywords Foggy image · Dark channel · Practicality · Guided filter

69.1 Introduction

In the computer vision application system, the exits of haze or fog greatly interferes with the normal work of the computer vision application system. Humans, vehicles, and other targets in the images are in the edge blur, color distortion and loss of contour information occurs. A lot of research has been done on haze removal in computer vision applications. The algorithm research of fog removal has made some achievements, and the technique based on the physical model can be separated into two classes according to whether the scene additional information is needed. One obtains more images under different states of weather [1–3], or varying degrees of polarization [4–7], to remove the fog, but these methods are unpractical. The other

Z. Wang (✉) · H. Li · J. Teng · X. He
School of Control and Computer Engineering, North China Electric Power University, 102206 Beijing, China
e-mail: cncepu@gmail.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_69

one estimates the model parameters and restores the images by trying to set out from the image itself, which is practical. Fattal [8] used a local window-based operation and a graphical model. However, it attempts to separate uncorrelated fields.

He [9]restored a fog-degraded image by using the dark channel prior and gained the sharp image; however, this method has a computational complexity and is time-consuming. In this chapter, we propose a novel type of image restoration, namely, the gray image guided filter. We find that the cause of fog-degraded images is single, and the degenerative characters of the gray image are in common with the degenerative characters of its color image; as a result, the transmission of the atmospheric scattering model estimates to use the dark channel prior, and optimizes the transmission by guiding a filter on the gray image to obtain a sharp image in this way. Our method can keep the edge of the input image to gain a sharp image. Experiments show that the improved method to get the same fog-free image is more efficient.

69.2 Defogging Based on Dark Channel Prior

69.2.1 Atmospheric Models

In computer graphics and computer vision, a very popular model to describe the formation of a fog-degraded image was proposed by W. E. K. Middleton [10] in 1952 with the expression shown as follows:

$$I(x) = J(x)t(x) + A(1-t(x)), \tag{69.1}$$

where $I(x)$ is the obtained image, $J(x)$ is the scene radiance, A is the atmospheric light, and $t(x)$ is the medium transmission describing the portion of the light that reaches the camera is not scattered. The aim of the fog removal is to separate J, A and t from I.

69.2.2 Dark Channel Prior

He proposed a simple but useful prior in his paper—dark channel prior. He found that in the vast majority of the local area of the sky, there always exists at least one color channel having a very low value at some pixels, which means the minimum intensity in such regions should have a very low value, converging to zero. As described in the formula, for a dark image $J(x)$, we define:

$$J^{dark}(x) = \min_{color \in \{r,g,b\}} \left(\min_{y \in \Omega(x)} (J^{color}(y)) \right) \tag{69.2}$$

Equation (69.2) is presented according to the definition of dark channel prior, and we call $J^{dark}(x)$ the dark channel of J. As to the fog-free image, after dark channel priority, most of the dark channel image pixel values are close to zero. As to the other way around, if there are a lot of high-value pixels in the dark channel image, these pixel values can represent the additional impact of the fog on the scene imaging, which is the airlight of the foggy weather. What should be noted is that we only consider the condition that the input image is captured on foggy days.

69.2.3 Fog Image Restoration

According to the dark channel prior, for the fog image, we can estimate the atmospheric light A and the transmission $t(x)$. In the transmission image, the spaces with higher pixel values will be less affected by the fog, namely, they are closer to the imaging device because we use block areas in the local area when calculating the dark channel image; in this sense, the transmission estimated through the dark channel image exits the block effect. For the better description of the image edge information, the global optimization of initial transmission is needed. In He's paper, he used soft matting to optimize the initial transmission; nevertheless, soft matting has a computational complexity which is time-consuming. The algorithm is as follows: First, hypothesize that atmospheric light A is given and the transitivity $\tilde{t}(x)$ in a local area is constant; then use the dark channel prior in Equation (69.1):

$$\min_{col}\left(\min_{y\in\Omega(x)}\left(\frac{I^{col}(y)}{A^{col}}\right)\right)=\tilde{t}(x)\min_{col}\left(\min_{y\in\Omega(x)}\left(\frac{J^{col}(y)}{A^{col}}\right)\right)+(1-\tilde{t}(x)), \tag{69.3}$$

where $\Omega(x)$ is the region centered at x, and col represents the three color channels.

According to the laws of the dark channel prior, the J^{dark} of the fog-free image should approach zero, then estimate the transmission $\tilde{t}(x)$.

He's paper used soft matting to optimize the initial transmission. Figure 69.1e is the result after soft matting, and as it shows, the optimized transmission can keep the sharp edge discontinuity and sketch the contours of the outline of the object.

The intensity of the sky in a fog image is always very similar to the atmospheric light, and can always been used to estimate the atmospheric light A. According to the atmospheric light A and the optimized transmission $\tilde{t}(x)$, we can obtain the fog-free image $J(x)$ by Eq. (69.1):

$$J_c(x)=\frac{I_c(x)-A_c}{\max(t(x),t_0)}+A_c. \tag{69.4}$$

Figure 69.1 shows the results according to the algorithm of image fog removal with an image size of 600×400, and an experimental condition as follows: Intel Pentium CPU, 2.13 GHz, RAM 2.0 GB, OS Windows 7. The result shows that without soft matting, (Fig. 69.1d) cannot get a good result for fog removal. Soft matting has been

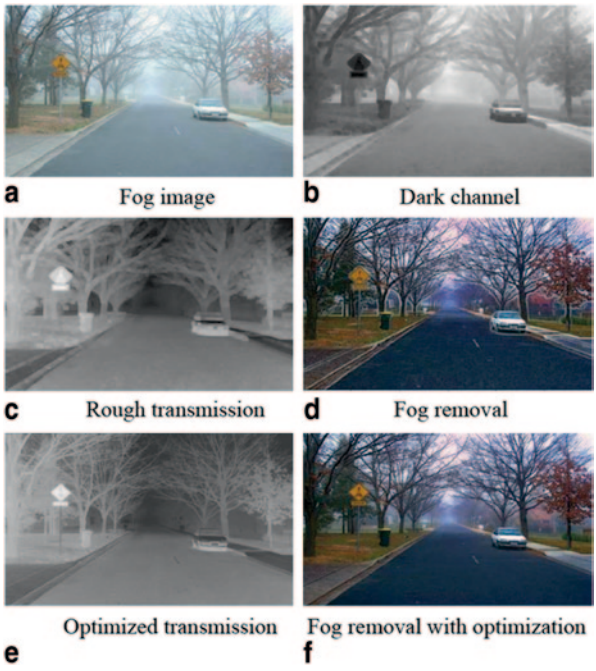


Fig. 69.1 Fog removal using dark channel prior. (**a**) Fog image, (**b**) dark channel, (**c**) rough transmission, (**d**) dark channel recover, (**e**) optimized transmission, (**f**) the results of using soft matting for optimization

used to optimize the algorithm (Fig. 69.1f) in removing the fog well and to keep the image information at the same time, but this method is time-consuming without effectiveness; as a result, in this chapter, we do not affect the subjective visual effect and sharply improve the processing speed of the algorithm.

69.3 Image Fog Removal Based on Gray Image Guided Filtering

As mentioned above, we can obtain a sharp image if the transmission is optimized. From the soft matting algorithm, the optimization mainly contains the regional smoothing and keeps the edge features; therefore, in this chapter, another algorithm to optimize the transmission is used. We use the gray image of the original image as the guide image and filter the transmission by structuring a variable coefficient filter. This filtering method is called guided filter [11].

69.3.1 Guided Filter

Guide filter assumption: The filter model is linear at the local area of the image, and in each local area the filtering output is a linear transformation of the guide image.

Assuming the guide image is $I(x)$, the image for filtering is $p(x)$, the filtered image is $q(x)$, we thus have:

$$q(x) = \sum_{j \in \omega_\kappa} W_j(I) p(j), \tag{69.5}$$

$$q(x) = \alpha_\kappa I(x) + \beta_\kappa, \forall x \in \omega_\kappa, \tag{69.6}$$

where ω_κ is a region centered at x and $W_j(I)$ is the variable filter core upon translation; it can be obtained by its guide image.

In order to solve the coefficients α_κ and β_κ, minimize the difference between q and p to solve the optimal solution by the definition above, thus we have:

$$E(\alpha_k, b_k) = \sum_{i \in \omega_k} ((\alpha_k I_i + b_k - p_i)^2 + \varepsilon \alpha_k^2). \tag{69.7}$$

In order to prevent α_κ from becoming too large, we set up the penalty factor ε. By solving the optimal solution of the secondary system above, α_κ and β_κ can be obtained as follows:

$$\alpha_k = \frac{\frac{1}{\omega} \sum_{i \in \omega_k} I_i p - \mu_k \overline{p}_{ki}}{\sigma_k^2 + \varepsilon}, \tag{69.8}$$

$$b_k = \overline{p}_k - \alpha_k \mu_k, \tag{69.9}$$

where ω is the number of pixels in the filter window, μ_κ and σ_k are the average and variance of I in ω_k, and $\overline{p}_\kappa$ is the mean of $p(x)$ in ω_k. Putting α_κ and β_κ into Eq. (69.6), we can gain the filtered image $q(x)$.

69.3.2 Fog Removal Based on Gray Image Guided Filter

First of all, use the dark channel method to estimate a rough transmission $\tilde{t}(x)$ and atmospheric light A, as described in Sect. 2. Then, use the guided filter to implement the regional smoothing and keep the edge features. Use $I_{gray}(x)$ of $I(x)$ as the guide image to guide the $\tilde{t}(x)$ filtering. α_κ and β_κ can be gained by Eq. (69.8) and Eq. (69.9); then, puting α_κ and β_κ into Eq. (69.6), we obtain the transmission $t(x)$. Now, by Eq. (69.4), the sharp image $J(x)$ is restored.

Equation (69.5) and Eq. (69.6) show that the guided filter method is good to keep the image edge information,(Fig. 69.2) which implements the edge features.

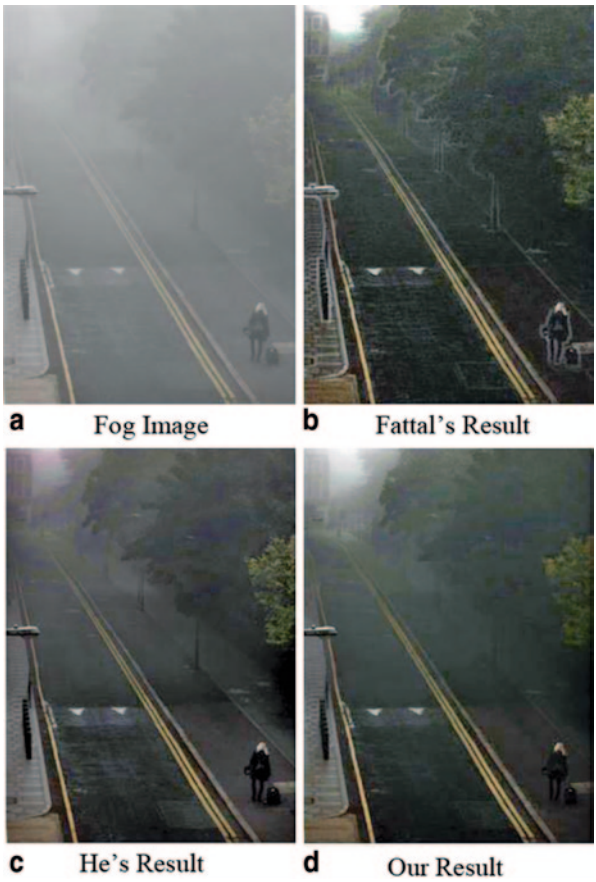


Fig. 69.2 Fog removal using dark channel prior. (**a**) Fog image, (**b**) dark channel, (**c**) rough transmission, (**d**) dark channel recover, (**e**) optimized transmission, (**f**) the results of using soft matting for optimization

69.4 Experimental Results

The experimental results show that, compared with Fattal's work, our algorithm has obvious improvement in terms of the visual effect; compared with He's work, our approach has obvious advantages in processing speed. The proposed approach can get a sharp image and improve the efficiency at the same time (Table 69.1).

Table 69.1 Processing time

	Name	Fig. 69.2
Size		346 × 472
Time (s)	Fattal's	0.247
	He's	3.226
	Gray image guided filtering	0.349

69.5 Conclusion

In this chapter, we use the gray image guide filter to optimize the transmission; the optimized transmission can realize the regional smoothing and keep the edge features. The experimental results show that the approach proposed in this chapter can optimize the transmission rapidly; therefore, our approach in fog removal maintains the image information including the edge of the image and the structure information; besides, we obtain a good visual effect. Objectively, the processing time is shortened greatly. The method proposed in this chapter is general about the 2D images. The dark channel prior assumes that the image exists in a dark area. If there are vast white areas in the foggy image, the dark channel prior would fail even though the guided filter is adopted. The following task will be to test whether it is suitable for 3D images and to explore ways to deal with the fog of large white areas.

Acknowledgements This work was supported in part by the Fundamental Research Funds for the Central Universities Grant.

References

1. Narasimhan SG, Nayar SK. Chromatic framework for vision in bad weather. Proceedings of IEEE conference on computer vision and pattern recognition. Washington D.C., USA: IEEE; 2000. pp. 598–605.
2. Nayar SK, Narasimhan SG. Vision in bad weather. Proceedings of the 7th IEEE international conference on computer vision. Kerkyra, Greece: IEEE; 1999. pp. 820–7.
3. Narasimhan SG, Nayar SK. Contrast restoration of weather degraded images. IEEE Trans Pattern Anal Mach Intell. 2003;25(6):713–24.
4. Schechner YY, Narasimhan SG, Nayar SK. Instant dehazing of images using polarization. Proceedings of the IEEE computer society conference on computer vision and pattern recognition. Washington D. C., USA: IEEE; 2001. pp. 325–32.
5. Namer E, Schechner YY. Advanced visibility improvement based on polarization filtered images. Proceedings of the polarization science and remote sensing II. San Diego, USA: SPIE; 2005. pp. 36–45
6. Schechner YY, Narasimhan SG, Nayar SK. Polarization-based vision through haze. Appl Opt. 2003;42(3):511–25.
7. Shwartz S, Namer E, Schechner YY. Blind haze separation. Proceedings of the IEEE computer society conference on computer vision and pattern recognition. Washington D. C., USA: IEEE; 2006. pp. 1984–91.
8. Fattal R. Single image dehazing. ACM Trans Graph. 2008;27(3):1–9.

9. HE KM, Sun J, Tang XO. Single image haze removal using dark channel prior. Proceedings of the IEEE conference on computer vision and pattern recognition miami. USA: IEEE; 2009. pp. 1956–63.
10. Middleton WEK. Vision through the atmosphere. Geophysics II. Berlin: Springer; 1957. pp. 254–87.
11. He K, Sun J, Tang X. Guided image filtering. Computer Vision–ECCV 2010. Berlin: Springer; 2010. pp. 1–14.

Chapter 70
Parallel Acquisition Architecture for DS_FH Hybrid Spread Spectrum Signal in TT&C

Hailong Deng and Wei Zhan

Abstract Parallel acquisition architecture with feedback loop in the time-frequency domain is proposed. After obtaining the coarse estimation of the carrier frequency offset by exploiting the partial match filter–fast Fourier transform (PMF-FFT) algorithm, this architecture uses the Quinn frequency interpolation algorithm to improve the estimation accuracy of the frequency offset and modifies the local carrier by employing the estimated result and reacquiring. Compared to the PMF-FFT algorithm, the new method improves the acquisition accuracy of the carrier and reduces the average acquisition time under the conditions of the large frequency offset and low signal-to-noise ratio (SNR).

Keywords DS_FH · Parallel acquisition · Acquisition accuracy · Mean acquisition time

70.1 Introduction

It is an attempt to introduce the DS_FH hybrid spread spectrum technology to spacecraft TT&C [1–3], but the challenge lies in the acquisition of the DS_FH hybrid spread spectrum signal under large carrier frequency offset. Supposing that the traditional acquisition method in time or frequency domain is applied without external assistance, the time needed is so long that it would severely constrain the hopping rate and reduce the antijamming performance of the DS_FH hybrid spread spectrum system [4]. The traditional partial match filter–fast Fourier transform PMF-FFT algorithm has a good property of fast acquisition, which can search carrier frequency and code phase in the time-frequency domain simultaneously [5, 6]. However, the frequency estimation accuracy of PMF-FFT algorithm exceeds the range of linear approximation of the nonlinear device in the carrier tracking loop

H. Deng (✉)
Logistics College of PLA, 100858 Beijing, China
e-mail: leshanleshui@163.com

W. Zhan
Army 71217 of PLA, 265200 Shandong, China
e-mail: 275600244@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_70

because of the FFT fence effect. This chapter proposes parallel acquisition architecture with feedback loop by combining the PMF-FFT algorithm and one frequency estimation algorithm which has the advantage of small amount of computation and real-time calculation. The new architecture uses the Quinn frequency interpolation algorithm to further estimate the carrier frequency offset and modifies the local carrier by employing the estimated result and reacquires. Under the circumstance of low signal-to-noise ratio (SNR) and large carrier frequency offset, this method can increase the estimation accuracy of the carrier frequency significantly and reduce the mean acquisition time.

70.2 Architecture Proposed

The analysis here will consider a double-dwell detector. Figure 70.1 shows the principle diagram. As it is shown, the first step is to obtain the coarse estimation of the frequency offset and spreading code phase using the PMF-FFT algorithm; the module Acquisition Judgment 1 works by judging coarse acquisition result, and then the Quinn frequency interpolation algorithm is exploited to further analyze the frequency parameter. Next, the PMF-FFT algorithm and module Acquisition Judgment 2 are exploited to obtain the carrier frequency offset when the local carrier frequency NCO (Numerically Controlled Oscillator) is adjusted by the refined estimation result.

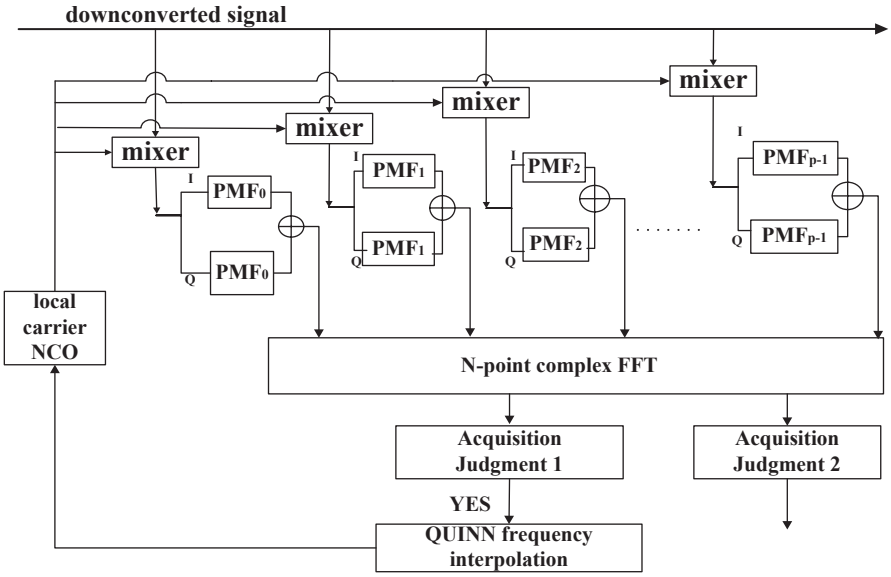


Fig. 70.1 Parallel acquisition architecture with feedback loop in the time-frequency domain

The normalized output of N-point FFT is given as Eq. (70.1)

$$\begin{aligned} S(k) &= \sum_{d=0}^{N-1} \frac{1}{2N} R(\Delta\tau) \frac{\sin(\pi f_{di} L T_s)}{L \sin(\pi f_{di} T)_s} \times e^{j[\pi(L-1) f_{di} T_s + \varphi_i]} e^{j 2\pi f_{di} L d T_s} e^{-j\frac{2\pi}{N} dk} \\ &= \frac{1}{2N} R(\Delta\tau) \frac{\sin(\pi f_{di} L T)_s}{L \sin(\pi f_{di} T_s)} e^{j[\pi(L-1) f_{di} T_s + \varphi_i]} \\ &\times \frac{\sin\left[\left(\pi f_{di} L T_s - \frac{\pi k}{N}\right) N\right]}{\sin\left(\pi f_{di} L T_s - \frac{\pi k}{N}\right)} e^{j\pi(f_{di} L T_s - \frac{k}{N})(N-1)}. \end{aligned} \tag{70.1}$$

In Eq. (70.1), L is the number of PMF, T_s is DS chip time, f_{di} is the carrier frequency offset between the transmitter and the receiver, and $R(\Delta\tau)$ is the autocorrelation function of the spreading code, where $\Delta\tau$ represents the spreading code phase difference in the transceiver.

Mark the discrete frequency index as k_0, which is the maximum value of s (k), $k_0 = \text{int}[f_{di} L T_s N]$; int [$x$] represents the integer closest to x. If N is a big number, s (k_0) can be approximated as follows:

$$S(k_0) = \frac{1}{2} R(\Delta\tau) \frac{\sin(\pi f_{di} L T_s)}{\sin(\pi f_{di} T_s)} e^{j[\pi(L-1) f_{di} T_s + \varphi_i]} \times \frac{N \sin(\pi\delta)}{\pi\delta} e^{j\pi\delta(N-1)/N}, \tag{70.2}$$

where δ is defined as the difference between the real carrier frequency and the estimated frequency, $\delta = f_{di} L T_s N - k_0$. According to the Quinn frequency interpolation algorithm, we can define

$$\alpha_1 = \Re\left\{S(\mathrm{k}_0 - 1) / S(\mathrm{k}_0)\right\} \tag{70.3}$$

$$\alpha_2 = \Re\left\{S(\mathrm{k}_0 + 1) / S(\mathrm{k}_0)\right\}, \tag{70.4}$$

where $\Re$ indicates the real part; substituting (70.2) into (70.3) (70.4), we can obtain

$$\alpha_1 = -\frac{\delta}{1+\delta} \cos\left(\pi \frac{N-1}{N}\right) \tag{70.5}$$

$$\alpha_2 = \frac{\delta}{1-\delta} \cos\left(-\pi \frac{N-1}{N}\right). \tag{70.6}$$

When N is a large number, we can simplify formulas (70.5) and (70.6) as

$$\alpha_1 = \frac{\delta}{1+\delta} \tag{70.7}$$

$$\alpha_2 = -\frac{\delta}{1-\delta}. \tag{70.8}$$

Converting the formula's form, (70.7) and (70.8) can be expressed as

$$\delta_1 = \alpha_1 / (1-\alpha_1) \tag{70.9}$$

$$\delta_2 = -\alpha_2 / (1-\alpha_2). \tag{70.10}$$

If δ_1 and δ_2 are bigger than zero, then $\delta = \delta_2$, else $\delta = \delta_1$. The result of frequency estimation is derived as

$$\hat{f} = (\mathrm{k}_0 + \delta) f_s / N. \tag{70.11}$$

70.3 Simulation Result

Compared with the PMF-FFT algorithm, the parallel acquisition architecture with feedback loop in the time-frequency domain improves the accuracy of carrier acquisition by using the Quinn frequency interpolation algorithm. Equation (70.12) presents the frequency estimation variance of the Quinn interpolation algorithm [7], which is also the acquisition accuracy of our method.

$$\delta_f^2 = \frac{(1-|\delta|^2)[1-|\delta|^2+\delta^2]}{T^2 \bullet N \bullet S_{NR} \bullet \sin c^2(\delta)}, \tag{70.12}$$

where T is the sampling time of the data length.

According to the simulation result, carrier frequency estimation deviation of our method is greatly reduced compared with the PMF-FFT algorithm, as Fig. 70.2 exhibits.

Simulation parameters: frequency-hopping rate, 100 hop/s; the total number of hopping frequency, 7; hopping interval, 25 MHz; data rate, 10kbps; spreading code length, 1023; code rate, 10.23 Mbps; SNR, -15dB; sampling rate, 512 Mbps; sampling points in one PMF, 400; the number of PMF, 128; make a 128-point FFT, doing Monte Carlo simulation 100 times in one frequency.

As the simulation result presents, the maximum frequency estimation deviation of the PMF-FFT algorithm is 5000 Hz, whereas that of the algorithm proposed in this chapter is only 20 Hz. The frequency estimation results of our method can enter the linear approximation range of the carrier tracking loop without frequency pulling.

In the circumstance of low SNR and large frequency offset, the algorithm proposed in this chapter would reduce the mean acquisition time.

When there is a signal, PMF-FFT output approximately obeys the Rice distribution, the probability of detection

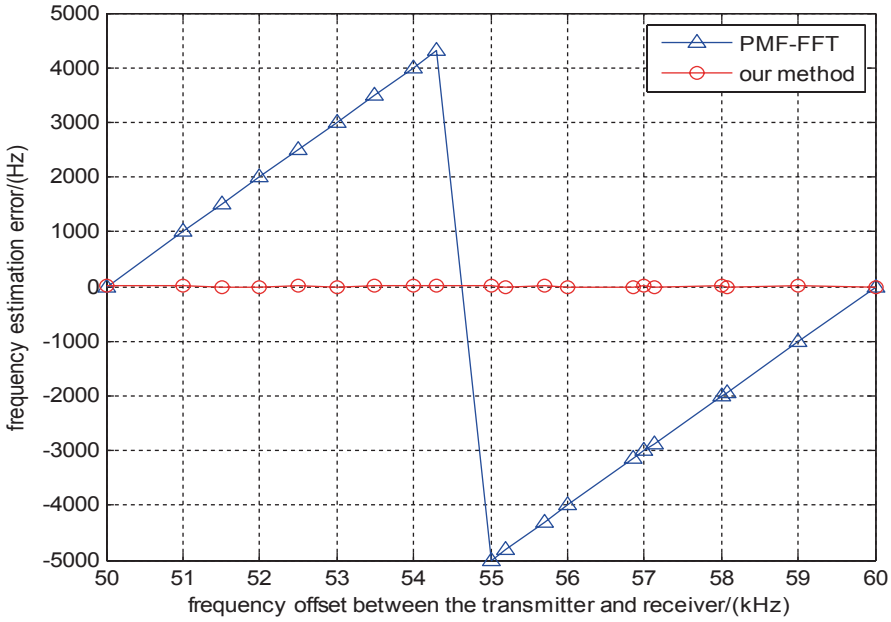


Fig. 70.2 Carrier frequency estimation error of the two algorithms

$$P_d = 1 - \prod_{k=1}^{N} (1 - P_k), \tag{70.13}$$

where $P_k = Q\left(\frac{s}{\sigma_n}, \frac{V_t}{\sigma_n}\right)$, Q is Marcum function, s is the mean energy of the signal, V_t is the decision threshold, and σ_n is MSE (Mean Squared Error) of noise.

When there is no signal, PMF-FFT output approximately obeys the Rayleigh distribution, the probability of false alarm

$$P_{fa} = 1 - \left(1 - \exp(-\frac{V_t^2}{2\sigma_n^2})\right)^N. \tag{70.14}$$

Under the double-dwell detection mode, when the awaiting search method is employed for hopping synchronization, the mean acquisition time of the two algorithms for the DS_FH hybrid spread spectrum signal is the same, as Eq. (70.15) shows:

$$\overline{T}_{acq} = \frac{2 + 2rP_{\mathrm{d1}} + (2 - P_{\mathrm{D}})(2 * M - 1)(KP_{\mathrm{FA}} + rP_{\mathrm{fa1}} + 1)}{4P_{\mathrm{D}}} \cdot T_c + \frac{(N_{\mathrm{h}} - 1)}{2} \cdot T_{\mathrm{h}}. \tag{70.15}$$

K is the penalty factor, R is the spreading code period for the verification stage, M is the length of the spreading code, N_{h} is the number of frequencies in the hopping pat-

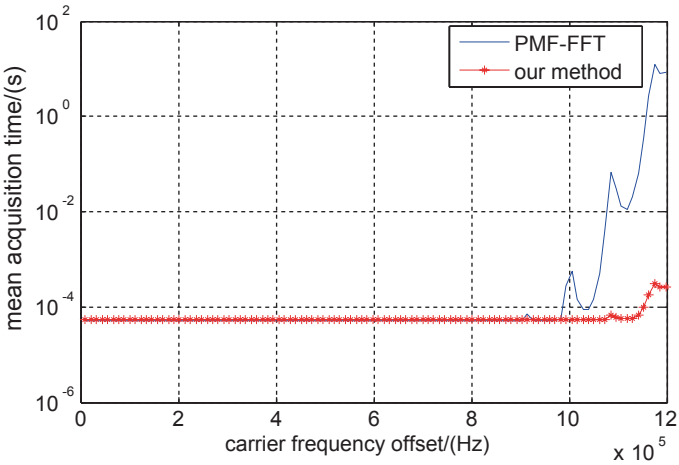


Fig. 70.3 The mean acquisition time of the two algorithms when dual-threshold is used

tern, T_h is the hopping frequency time, (P_{d1}, P_{fa1}) and (P_{d2}, P_{fa2}) are the probability of detection and probability of false alarm during the acquisition and verification stage, and $P_D = P_{d1} \times P_{d2}, P_{FA} = P_{fa1} \times P_{fa2}$.

As the awaiting search method is used by frequency-hopping synchronization, the average synchronization time of frequency-hopping synchronization has nothing to do with the direct-sequence spreading code's acquisition. In Eq. (70.15), the preceding part is the DS signal's average acquisition time, and the consequent part is the FH signal's average acquisition time.

When the dual-threshold is used by this chapter's method, the simulation result is shown in Fig. 70.3. Simulation parameters: Threshold in Acquisition Judgment 1 is 60; Acquisition Judgment 2's threshold is 200, other simulation parameters are the same as in Fig. 70.2.

As Fig. 70.3 exhibits, the mean acquisition time of this chapter's method is far less than the PMF-FFT algorithm. For example, under the carrier frequency offset of 1.08e5Hz, the mean acquisition of our method is 1e-4 s, while the mean acquisition of the PMF-FFT algorithm is 1e-2 s. That is because our method can greatly reduce the energy loss caused by incomplete phase compensation of FFT, thereby increasing the detection probability of acquisition and reducing false alarm probability. According to Eq. (70.15), a larger detection probability or a smaller probability of false alarm can get smaller average acquisition time.

70.4 Conclusion

In this chapter, the parallel acquisition architecture in the time-frequency domain was proposed. In contrast to the PMF-FFT algorithm, the new method increased the acquisition accuracy of carrier frequency, which was estimated so well that it can

pull the frequency estimation result into the linear approximate range of the phase detector in the carrier tracking loop. Under the circumstance of low SNR and large carrier frequency offset, the new method could reduce the mean acquisition time significantly. It is a good method for fast synchronization of spread spectrum signal. However, the Quinn frequency interpolation algorithm utilizes two lines around the peak of the spectrum to estimate frequency; when zero is padded to the FFT result, the frequency estimation error will increase sharply because the line value of the spectrum has changed. Therefore, the new method is not suitable for the zero-padded FFT.

References

1. Lorenzo S, et al. Frequency-hopping techniques for secure satellite TT&C: System analysis & trade-offs. Alcatel Alenia Space-Italy, IEEE. 2006;12(3):55–60.
2. Wang C, Nguyen TM, et al. On the power spectral density of SGLS and USB waveforms. 1999 IEEE Aerospace Conference, IEEE; 1999. p. 137–43.
3. Xiao K, et al. An uplink carrier frequency offset tracking method in aeronautical communications system. International workshop on high mobility wireless communications, IEEE; 2013. p.14–8.
4. Lou D, et al. The range and velocity measurement performance of DS/FH hybrid spread spectrum signal with the interference. 2nd international conference on computer science and network technology, IEEE; 2012. p. 1405–8.
5. Liu C, Zhang J. Analysis and optimization of PMF-FFT acquisition algorithm for high-dynamic GPS signal. 2011 IEEE 5th international conference on cybernetics and intelligent systems, IEEE; 2011. p. 185–9.
6. Qi H, Shi X, Ji L. PMF-FFT Algorithm For Pn code acquisition. J Xi'an Technol Univ. 2010;30(1):57–61.
7. Qi G. Precision analysis of several fft-based frequency estimation methods. J Vib Eng. 2006;19(1):86–92 (In Chinese).

Chapter 71
Monitoring CE'3 Rover Movement Using Same-Beam Interferometry with China's Deep Space Network

Songtao Han, Geishi Tang, Jianfeng Cao, Lue Chen and Tianpeng Ren

Abstract China's Deep Space Network (DSN) was deployed in 2013, a high-accuracy interferometric tracking observable was successfully used for orbit determination in the CE'3 project. This chapter presents the results of the rover's status monitoring by earth-based same-beam interferometry (SBI) measurement. Phase delay with a biased offset is induced to identify the rover's movement status, which could be up to three orders of magnitude better than conventional group delay. Based on the tracking data within China's DSN, the status of the rover including going ahead, turning over, and changing signal can be successfully identified with a higher resolution. Combined with baseline length within China's DSN, the movement of the rover in the order of centimeters can be identified, which testifies the effectiveness of this technique and algorithm.

Keywords CE'3 · Phase delay · Rover · SBI · VLBI

71.1 Introduction

As a key milestone of the China lunar project, CE'3 was launched on December 2, 2013. After cruising to the moon for about 12 days, CE'3 made a soft landing at the plains of Sinus Iridum on the lunar surface. The rover departed from the lander and moved around the lander to perform scientific experiments. Vision-based navigation is used for the rover tracking and location. Pictures taken by the navigation camera are used to match a database and then the location of the rover can be identified. Because the picture can only be taken when the rover is static, the process of the movement cannot be monitored. In order to support deep space navigation, China's Deep Space Networks (China's DSN) were originally employed in 2013. The Joint Space Operations Center Mission System (JMS) tracking station with an antenna dish of 66 m and KSH tracking station with an antenna dish of 35 m were

S. Han (✉) · G. Tang · J. Cao · L. Chen · T. Ren
National Key Laboratory of Science and Technology on Aerospace Flight Dynamics, Beijing Aerospace Control Center, 100094 Beijing, China
e-mail: justdoit_doing@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_71

firstly applied in the CE'3 project; Delta-differential one-way range (DOR) measurement was successfully taken for the orbit determination of CE'3. In this chapter, we focus on the rover's movement tracking using the same-beam interferometry (SBI) measurement by China's DSN.

The theory of SBI measurement is described briefly, then the algorithm is discussed. The results of tracking CE'3 rover are presented in Sect. 4; also performance is analyzed.

71.2 Theory of SBI Tracking

Earth-based radiometric tracking is widely used in spacecraft navigation, such as ranging and Doppler tracking [1, 2]. Very long baseline interferometry (VLBI) was firstly induced to spacecraft navigation in the 1960s, and it is used to determine the target angular position [3, 4]. As a working mode of VLBI, SBI can measure the angular separation of the two spacecraft in the plane of sky [5]. When the separation angle is smaller than the antenna beam width, the two targets can be viewed by the antenna simultaneously; most of the observable errors will be cancelled by differential processing [6, 7]. In general, SBI measurement accuracy could be in the order of several picoseconds in the X band [8]. Figure 71.1 shows the tracking geometry of SBI.

Delay observable by the correlator can be expressed as:

$$\begin{aligned} \tau_A &= \tau_{g\_A} + \tau_{A\_clc} + \tau_{A\_\mathrm{med}} + \tau_{A\_\mathrm{path}} \\ \tau_B &= \tau_{g\_B} + \tau_{B\_clc} + \tau_{B\_\mathrm{med}} + \tau_{B\_\mathrm{path}}, \end{aligned} \tag{71.1}$$

where subindex A, B correspond to the two targets, respectively, and τ_{g_A} presents the geometry delay caused by the triangular combined with the baseline and the

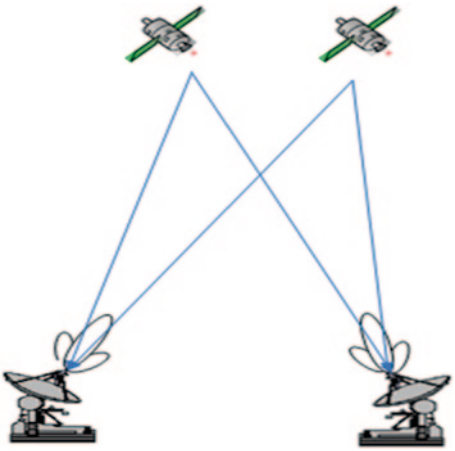


Fig. 71.1 Tracking geometry of SBI

target, τ_{A_clc} is the station synchronization, τ_{A_med} is the media transverse delay, and τ_{A_path} is the instrumental delay.

In theory, the differential delay of the two targets only depends on the geometry, and it can exactly reflect the relative position between the two targets in the plane of sky. The accuracy of SBI measurement depends on the following factors [9, 10].

The media delay depends on the pointing direction of the antenna. It usually includes two kinds of media effects. The error source included by the ionosphere can be expressed as:

$$\varepsilon_d = \sqrt{2} \bullet \Delta\theta \bullet F \bullet \delta, \tag{71.2}$$

where $\Delta\theta$ is the separation angle between the two targets, F is the derivative of the mapping function in the direction of $\Delta\theta$, for X band, and $\delta \approx 0.03m$.

The delay effect induced by troposphere is:

$$\varepsilon_d = \sqrt{2} \bullet \Delta\theta \bullet 0.04 \bullet \frac{\cos E}{\sin^2 E}, \tag{71.3}$$

where E is the elevation angle.

Another error source for SBI measurement is a system random error, which can be expressed as:

$$\varepsilon_d = \sqrt{2}\,\frac{\lambda}{2\pi \mathrm{SNR}}, \tag{71.4}$$

where λ is the wavelength and *SNR* is the signal-to-noise ratio:

$$SNR = \sqrt{\frac{4P}{\pi N_0} T}, \tag{71.5}$$

where P is the receiving power of target, N_0 is the system noise within 1 Hz bandwidth, and T is the integration time.

71.3 Processing Analysis

Processing algorithm of acquiring interferometric phase is the same as traditional VLBI, including bit shift compensation, fringe stopping, and fractional bit correction [11].

We discuss the postprocessing algorithm. Interferometric phase for the two targets φ_A, φ_B can be expressed as:

$$\varphi_A + N_A \bullet 2\pi = 2\pi f_A \bullet \tau_A \tag{71.6}$$

$$\varphi_B + N_B \bullet 2\pi = 2\pi f_B \bullet \tau_B, \tag{71.7}$$

where τ_A, τ_B are the delay of the two targets separately, N_A, N_B are the phase blur numbers, and f_A, f_B are the carrier frequencies.

71.3.1 Phase Delay

Usually, VLBI computes group delay with bandwidth synthesis; the group delay accuracy depends on both interferometric fringe accuracy and the spanned bandwidth [12, 13]:

$$\tau = \frac{\phi}{\Delta f}. \tag{71.8}$$

If we substitute the synthesis bandwidth Δf with the signal frequency f, then we can get phase delay:

$$\tau = \frac{\phi}{f}. \tag{71.9}$$

The derivative of the above formulation is:

$$\begin{cases} \delta\tau = -\dfrac{\phi}{f^2}\delta f \\ \delta\tau = -\dfrac{\phi}{f}\delta\phi \end{cases}. \tag{71.10}$$

Frequency f is much larger than synthesis bandwidth, then accuracy of phase delay can be improved up to 2–3 orders of magnitude than group delay.

71.3.2 Phase Unwrapping

One of the key techniques is phase unwrapping to compute phase delay. Many methods are discussed in the literatures, such as group-delay-assisted method, combining multifrequency signal method [14].

In this chapter, we focus on the relative movement of the rover to the lander, which means the variety of the delay is the observable we care about, not the absolute delay. So a continuous interferometric phase with a constant offset can be used for monitoring the rover's movement.

In order to get a continuous interferometric phase with time variable, a prior model $\xi(\tau)$ used for correlation must satisfy the following formulation:

$$\left|\xi(\tau)_T - \xi(\tau)_0\right| < \frac{1}{f}. \tag{71.11}$$

The above formulation means the accuracy of prior delay rate model within integration time T must be lower than reciprocal of the signal frequency.

71.4 Experimental Results

CE'3 made a successfully soft landing at the plains of Sinus Iridum on the lunar surface on December 14, 2013. Then, the rover drove from the lander automatically. Since then, the rover has moved more than 100 m until the end of its lifetime.

China's DSN takes responsibility of monitoring the rover's movement. Two strategies are adopted: the first is vision navigation and the second one is radiometric tracking. Figure 71.2 shows the total motion trajectory of the rover.

The lander is equipped with an X-band transponder, which can transmit X-band DOR signal to the ground. Both DOR signal and wideband data communication are supported by the lander, only data communication signal is supported by the rover. There are totally four DOR tones coherent to the main carrier. The maximum spanned bandwidth of DOR signal is about 38 MHz [15, 16].

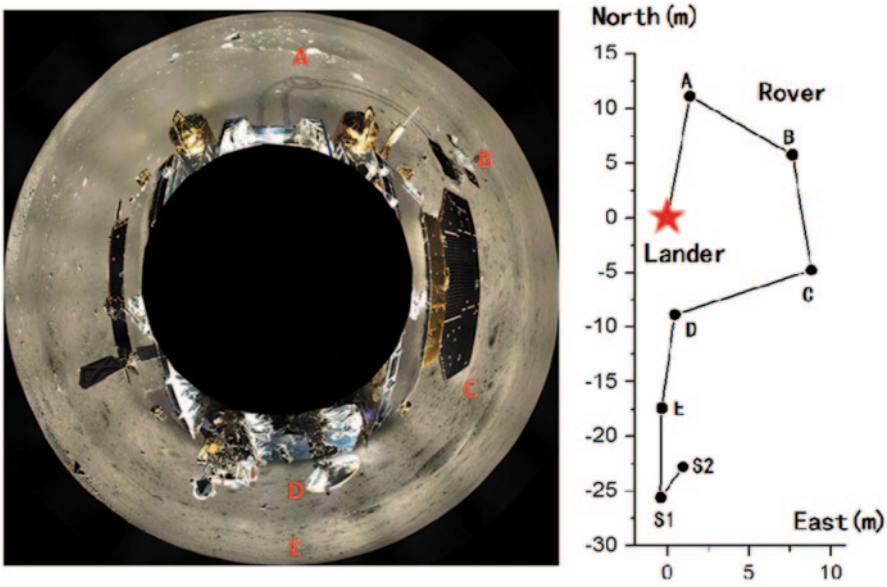


Fig. 71.2 Motion trajectory of the rover

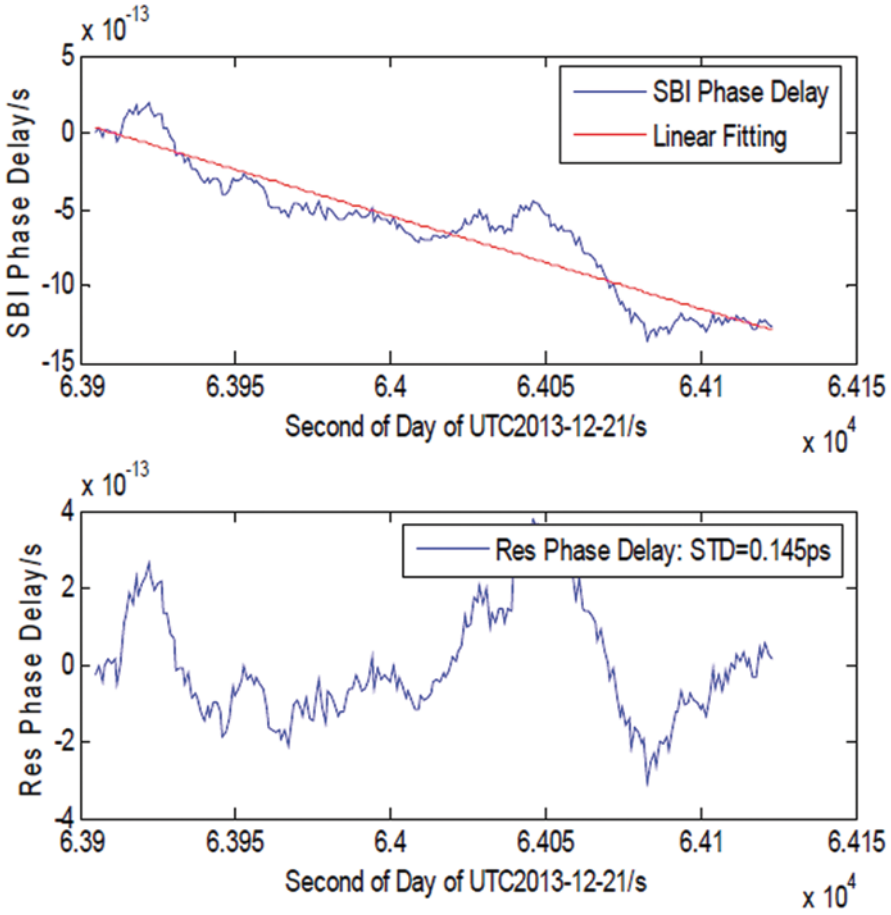


Fig. 71.3 Monitoring rover's go ahead

71.4.1 Rover's Going Ahead

Because the navigation camera cannot take pictures during the movement, monitoring only depends on radiometric tracking.

The accuracy of the observable is in the order of sub-ps, which corresponds to the uncertainty of centimeter in the plane of sky (Fig.71.3).

71.4.2 Rover's Turnover

The trajectory line of the rover is like an eclipse, the rover made a turnover when it changed its moving direction, and the interferometric phase exactly reflected the whole course, as shown in Fig 71.4.

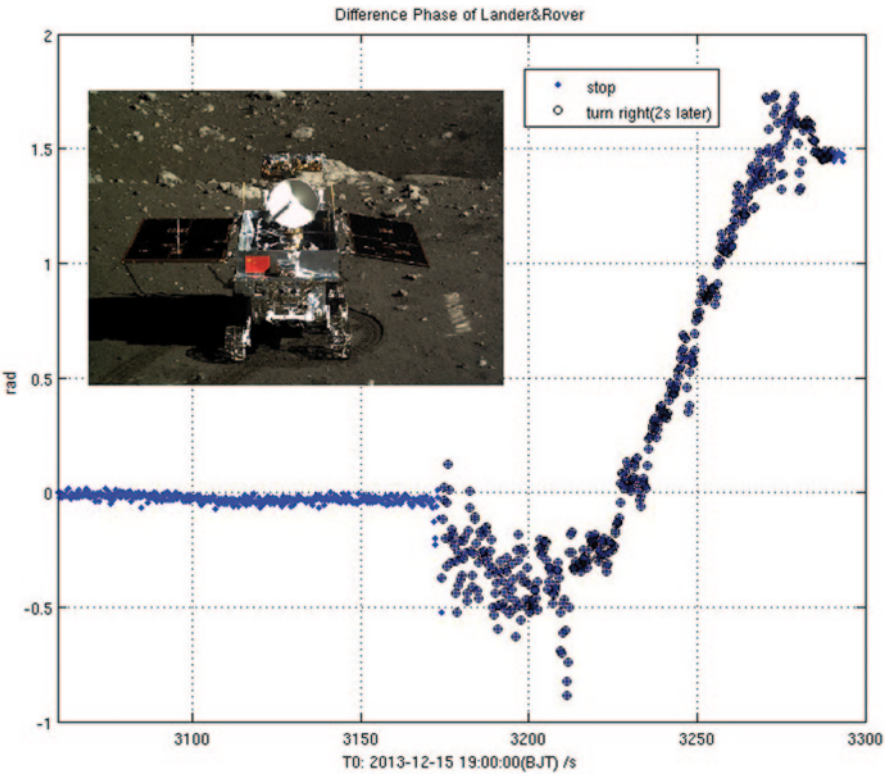


Fig. 71.4 Monitoring rover's turnover

71.4.3 Rover's Changing Signal

The data communication signal of the rover is transmitted by directional antenna, while telemetry signal is transmitted by the omnidirectional antenna. The two antennas are mounted in different positions, so in theory a gap exists in the SBI delay observable (Fig. 71.5).

The magnitude of the gap is about 45 ps, which corresponds to about 1.6 m in the plane of sky, which has the same size in the order of meters of the rover.

71.5 Conclusion

SBI uses two separated antennas to take a measurement of the two spacecrafts in the plane of sky. Because phase delay is induced, it can be predicted that the accuracy could be improved up to three orders than conventional group delay. In this chapter, SBI measurements for monitoring CE'3 rover movement are discussed, and the rover's movement at the centimeter level can be identified.

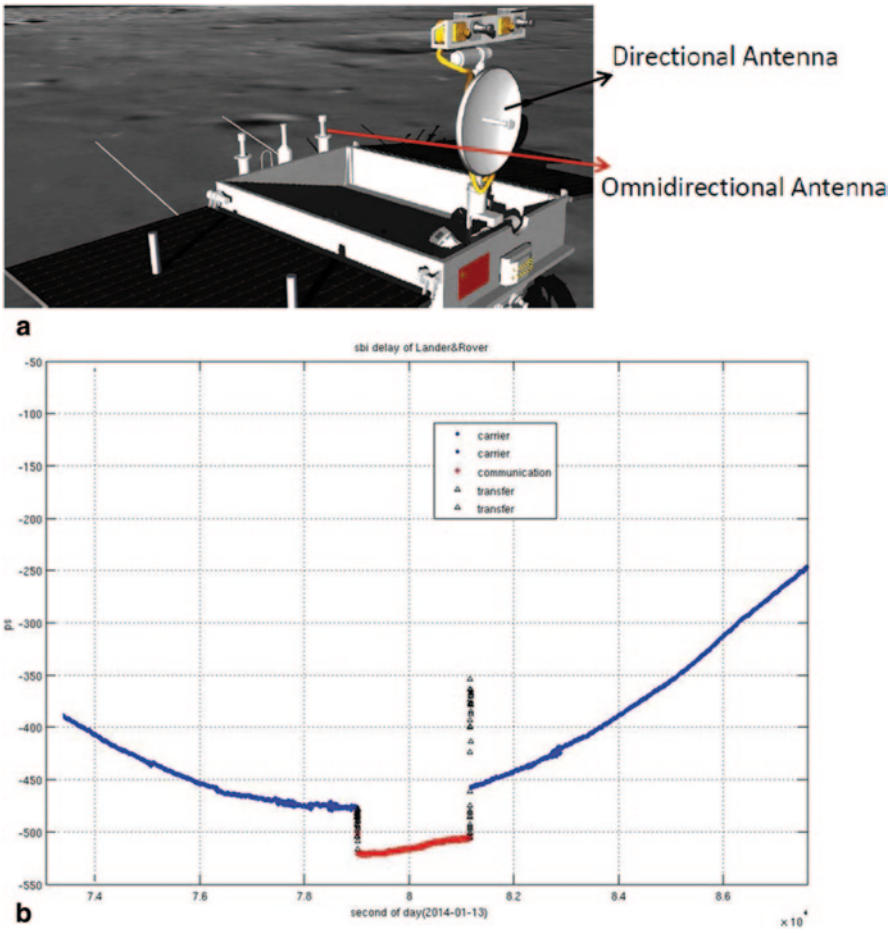


Fig. 71.5 Monitoring rover's changing signal. **a** Antenna displacement of the rover. **b** Observable for different signal

References

1. Yu ZJ. Deep space TT & C system. 1st ed. Beijing: National Defense Industry Press; 2009. p. 122–5.
2. Zhu XY, Li CL, Zhang HB. A survey of VLBI technique for deep space exploration and trend in China current situation and development. J Astron. 2010;31(8):1893–9. (In Chinese)
3. Tang GS. Radiometric measurement technique for deep space navigation. 1st ed. Beijing: National Defense Industry Press; 2012. p. 81–99.
4. Wu WR, Wang GL, Jie DG. High-accuracy VLBI technique using ΔDOR signals. Sci China Inform. 2013;43(2):185–96. (In Chinese).
5. Salzberg IM. Tracking the Apollo Lunar Rover with interferometry techniques. Proceedings of IEEE; 1973. p. 1233–6.

6. Preston R, Hildebrand C, Purcell JG. Determination of venus winds by ground-based radio tracking of the VEGA balloons. Science. 1986;231(4744):1414–6.
7. Kono Y, Hanada H, Ping JS, et al. Precise positioning of spacecraft by multi-frequency VLBI. Earth Planets Space. 2003;55:581–9.
8. Liu QH, Shi X, Wang WH, et al. High accuracy determination of very long baseline interferometry phase delays and its applications in China's future missions. Physics. 2009;38(10):712–6. (In Chinese).
9. Folkner WM, Border JS. Orbiter-orbiter and orbiter-lander tracking using same-beam interferometry. TDA progress report; 1992. 42–109:74–86.
10. Chen SW, Dong GL, Li HT, et al. Analysis on the impact of error on performance of real-time ambiguity resolving of SBI. Proceedings of the Second International Conference on Instrumentation, Measurement, Computer, Communication and Control; IEEE; 2012. p. 1591–5.
11. Deller AT, Tingay SJ, Bailes M. DiFX: A software correlator for very long baseline interferometry using multiprocessor computing environments. Publ Astron Soc Pac. 2007;119(853):318–36.
12. Lanyi G, Bagri DS, Border JS. Angular position determination of spacecraft by radio interferometry. Proceedings of IEEE; 2007. 95:2193–201.
13. Chen M, Liu QH. Study on differential phase delay closure of same-beam VLBI. 2nd International Conference on Computer Engineering and Technology; IEEE; 2010. p. V5-88–V5-92.
14. Liu QH, Chen M, Xiong WM, et al. Lunar rover location based on high accuracy multi-frequency SBI technique. Sci China Phys, Mech & Astron. 2010;40(2):253–60.
15. Wei EH, Jin SG, Yang HZ, et al. Simulation and results on real-time positioning of Chang'E-3 rover with the same-beam VLBI observations. Planet Space Sci. 2013;84:20–7.
16. Huang Y, Hu XG, Liu QH. Relative positon determination between lunar lander and rover using same beam VLBI technique. 44th lunar and planetary science conference; CLEP & Chinese academy of sciences; 2013. p. 1083–4.

Chapter 72
Color Image Segmentation Using Multilevel Thresholding—Hybrid Particle Swarm Optimization

Yang Liu, Kunyuan Hu, Yunlong Zhu and Hanning Chen

Abstract This chapter presents a novel multilevel threshold approach based on improved particle swarm optimization called as hybrid particle swarm optimization (HPSO), which can combine the advantages of the particle swarm optimization and the neighbor searching of artificial bee colony algorithm for color image segmentation. Experimental results show that the proposed approach adopts the multilevel threshold technique based on the improved HPSO algorithm to obtain a higher quality adequate segmentation; at the same time, it can also reduce the CPU processing time and eliminate the particles falling into local minima.

Keywords Image segmentation · Multilevel thresholding · Particle swarm optimization (PSO) · Hybrid particle swarm optimization (HPSO)

72.1 Introduction

The color image segmentation is the critical and essential step in the field of image processing, pattern recognition, and artificial intelligence. It can affect the final quality of the image processing. Its main task is divided into different zones according to the special meaning [1, 2, 4]. Compared to the gray images, the color images include luminance, hue, and saturation, so the segmentation and analysis in this regard are a challenge in the image processing arena [3]. The existing image segmentation approach generally includes the following categories: region growing edge detection, histogram thresholding, neural networks, clustering, and fuzzy logic [4]. The thresholding-based method is one of the widely used techniques. According to the selected threshold, the image can be divided into related sections. And

Y. Liu (✉) · K. Hu · Y. Zhu · H. Chen
Shenyang Institute of Automation Chinese Academy of Sciences, 110016 Shenyang, China
e-mail: liuyang4@sia.cn

Y. Liu
Shenyang University, 110044 Shenyang, China

University of Chinese Academy of Sciences, 100049 Beijing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_72

how to efficiently select the optimal thresholding is the key of image segmentation. Many methods for segmentation have been proposed in these years [5]. It introduces an improved color image segmentation algorithm based on hybrid particle swarm optimization (HPSO) to select the threshold, so the process is regarded as an optimization problem through this approach. In order to improve the traditional PSO algorithm, a domain search strategy is used, with the results showing that the proposed approach can execute faster, can eliminate the particles from falling into local minima, and is more stable than the traditional PSO algorithm [6].

72.2 Color Image Segmentation

72.2.1 Conversion of Color Space

The traditional color image segmentation approach can be regarded as an extension of the gray images. The color images can be seen as the possible combinations of three primary color components: red, green, and blue. Because the R, G, and B components have high correlation, different color representations can be derived from either linear or nonlinear transformations of RGB space, such as YIQ, YUV, HSV, and CIE.

HSV is a generally used image processing system. H represents hue, S represents saturation, and V represents value [4]. The pattern of HSV can be extended to color images, and the impact strength of vectors of the color information can be eliminated [4]. We can convert RGB image to HSV space, where H represents different colors, S represents the color shades, and V represents darkness or lightness. RGB space converted to HSV space formula is defined as follows:

$$S = 1 - 3/(R+B+G)[\min(R,G,B)]. \tag{72.1}$$

If $R \neq B$ *or* $R \neq G$

$$H = \cos^{-1}\left\{\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}. \tag{72.2}$$

If $B > G$

$$H = 2\pi - H \tag{72.3}$$

$$V = (R+B+G)/3. \tag{72.4}$$

Each component of a color space can adopt the proposed approach, and then the final segmentation results can be combined in some way.

72.2.2 The Multilevel Threshold

The image segmentation problem can be configured as multilevel optimization problem. When several distinct objects are contained in the image, the proper multiple threshold values should be selected for segmentation [1, 4].

Let there be L intensity levels [0, 1…L−1] of a given image; then,

$$p_i = \frac{h_i}{N}; \sum_{i=1}^{N} p_i = 1, \tag{72.5}$$

where the specific intensity level is represented by i, h_i represents the number of pixels for the corresponding intensity level I, and the total number of pixels is represented by N in the given image; thus, the total mean μ_T can be calculated as follows:

$$\mu_T = \sum_{i=1}^{L} ip_i = 1, \tag{72.6}$$

Assuming t_j is the generic n-level thresholding (j =1…n−1), and divide n classes of the total number of pixels to $D_1 \ldots D_n$, then the maximum between-class variance can be calculated as follows:

$$\sigma_B^2 = \sum_{j=1}^{n} w_j (\mu_j - \mu_T)^2 \tag{72.7}$$

$$w_j = \sum_{i=1}^{L} p_i, 1 \le j \le n \tag{72.8}$$

$$\mu_j = \sum\nolimits_{i=1}^{L} \frac{i_{p_i}}{w_j}, 1 \le j \le n, \tag{72.9}$$

where the specific class in such a way is represented by j, the probability of occurrence is represented by w_j, and the mean of class j is represented by μ_j.

We can convert the problem of n-level thresholding into an optimization problem to search the thresholds t_j; it can correspond to the maximized objective functions, which is generally defined as follows:

$$\phi = \max \sigma_B^2(t_j). \tag{72.10}$$

The stability of the algorithm is evaluated by the standard deviation, which can be calculated as follows:

$$STD = \sqrt{\sum_{i=1}^{n} \frac{(\sigma_i - \mu)^2}{N}}, \tag{72.11}$$

where N is the repeated times of each algorithm, the best fitness value of the i-th run of the proposed algorithm is represented by σ_i, and the average value of σ_i is represented by μ [6].

72.3 Color Image Multithreshold Segmentation Based on HPSO

In this chapter, a HPSO combines the merits of particle swarm optimization (PSO) and the neighbor searching of artificial bee colony algorithm while the domain search strategy is used to find the best position of the particle so as to solve the problem of particles falling into local minima [1, 7].

72.3.1 Traditional Particle Swarm Optimization

The traditional PSO model is the potential solution. The particle in the search space can be seen as a bird. Every particle has its own position and velocity. Empirical values are associated with the two elements which is not only its previous best position but also the best position of its neighbors. The swarm size is usually kept constant through iterations. Every particle's direction of movement can be calculated as follows:

$$v_i^{t+1} = x(v_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest_i - x_i^t)) \tag{72.12}$$

$$x_i^{t+1} = x_i^t + v_i^{t+1}. \tag{72.13}$$

In mathematical terms, v_i represents the rate of velocity for particles, $v_i = (v_{i1}, v_{i2}, \ldots v_{iD})$; x_i represents the i-th particle in the D-dimensional space, $x_i = (x_{i1}, x_{i2}, \ldots x_{iD})$; *pbest* is known as the best position of the i-th particle, *gbest* is known as the best position of any particles in its neighborhood. The constriction coefficient is represented as $\varkappa$, the learning rates are known as c_1 and c_2, two random vectors are known as r_1 and r_2, which are uniformly distributed in [0, 1], and t represents the time step [6, 7].

72.3.2 Hybrid Particle Swarm Optimization

According to formula (72.13), we should search the population particle for each particle and regard the searching result as the domain particles. While searching the entire field, if individual best of the domain particle is better than the historical individual best position, let the domain best be the local best of the particle.

$$new_{xij} = x_{ij} + r(x_{ij} - x_{kj}) \tag{72.14}$$

We choose indexes of k and j randomly, where k is distributed in [1, 2,…size] and j is distributed in [1, 2,…, D], and the values of k and i are not the same, r is chosen by the random number between [−1,1].

With the increase in the number of iterations, the value of $(x_{ij} - x_{kj})$ will gradually decrease. Another explanation is that the searching space of the particle will gradually shrink, which will also help improve the accuracy of the algorithm.

In the multidimensional space, the position and velocity values of every particle not only highly depend on the information of the local best but also depend on information of the domain best and the global best [8].

72.3.3 Color Image Multithreshold Segmentation Based on HPSO

The RGB image is converted into HSV space, where different colors are represented by *H,* the density of color is represented by *S*, and the luminance of color is represented by *I*. The relevance of each component of the color image is eliminated, and every vector of *H, S, V* adopts the multithreshold segmentation method based on HPSO [3]; thus, the result of the final image segmentation combines the three vectors. Figure 72.1 illustrates the flow chart of the algorithm [1, 3].

The fitness function of the particles in the HPSO is known as the between-class variance σ_B^2 of the image intensity distributions which can be calculated by formula (72.10).

The steps of the segmentation procedure are described in Table 72.1 [1]:

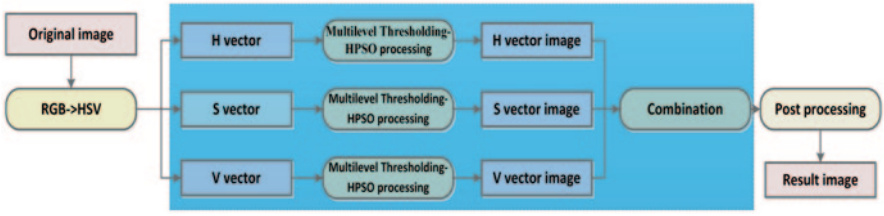


Fig. 72.1 Color image multithreshold segmentation based on HPSO

Table 72.1 Main steps of the segmentation procedure

Step 1: Convert the RGB image to HSV space
Step 2: Get the corresponding segmentation image at the three vectors by using the proposed algorithm
a. Initialize the swarm: Initialize the particles' positions and velocities, set the particles' velocities to zero, set the position randomly within the boundaries of the search space, the number of intensity levels L decides the search space, deploys the particles between 0 and 255 if the frames are 8-bit images
b. Update the particles' fitness according to formula (72.8) for each particle in the swarm
c. Evaluate the domain best by the fitness function. If the personality best position of the domain particle is better than the historical best of the particle, let the domain best be the local best of the particle
d. Update the particles' velocity and position vectors according to formulas (72.12) and (72.13)
e. Predefine the number of iterations of the stopping criteria, and the global best is the optimal threshold; otherwise go to step b
f. The optimal threshold is used for three vectors of the image segmentation
Step 3. Combine the three vector results, and then get the finial segmentation image

72.4 Experimental Result

In this section, the traditional PSO algorithm and the HPSO algorithm were programmed in MATLAB [1]. The population size for the two methods are set to 30, the maximum iterations of the two methods are set to 100, the learning rates are $c_1 = c_2 = 2.0$, and a total of 20 runs are performed [9]. The image of "peppers" is tested to segment the image for 2, 3, 4 thresholds with the PSO algorithm and the HPSO method. From Figs. 72.2 and 72.3, it can be seen that the adequate segmented result can be obtained by using the HPSO algorithm.

The standard deviation and CPU processing times are given in Table 72.2; the standard deviation values of the HPSO algorithm are more stable than the traditional PSO, which requires less CPU processing time for finding thresholds, especially in regard to higher threshold numbers. The results illustrate that the HPSO algorithm is more efficient than the traditional PSO.

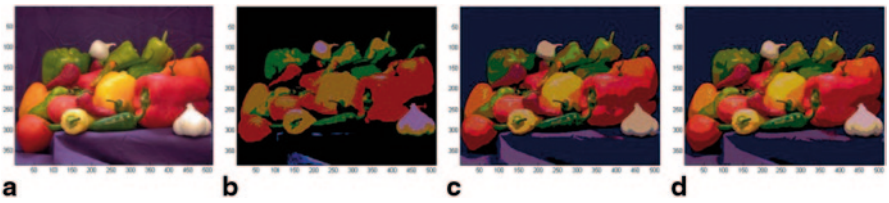


Fig. 72.2 The 2-, 3-, 4-threshold segmented result with PSO of "peppers." **a** Original. **b, c,** and **d** The 2-, 3-, 4-threshold segmented result

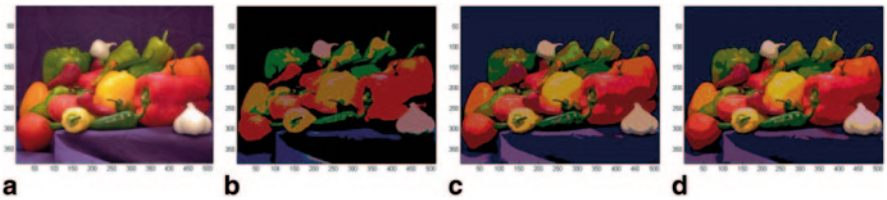


Fig. 72.3 The 2- , 3-, 4-threshold segmented result with HPSO of "peppers."**a** Original. **b, c,** and **d** The 2-, 3-, 4-threshold segmented result

Table 72.2 STD and CPU process time of different methods

	Thresholds	PSO	HPSO
STD	2	0.0790	0.0091
	3	1.0938	0.1059
	4	2.7456	0.2734
Computational time	2	0.3741	0.3465
	3	0.4967	0.4376
	4	0.5635	0.4879

72.5 Conclusion

In this chapter, a novel multilevel threshold method based on HPSO is proposed, which can overcome the disadvantages of the traditional PSO in local optimum conditions, and combines the advantages of PSO and the neighbor searching of artificial bee colony algorithm for every component of the color image space. The final segmentation results can be obtained by combination in some way. The results indicate that the proposed algorithm is more efficient than the traditional PSO.

Acknowledgments This work was supported in part by the National Natural Science Foundation of China under Grant 61105067 and 61174164, the General Project of Education Department of Liaoning Province under Grant L2013446, the Engineering research center of the Information Technology Integration (IOT) of Liaoning Province open-funded projects, and Key Laboratory of Networked Control System (CAS).

References

1. Ghamisi P. An efficient method for segmentation of images based on fractional calculus and natural selection. Expert Syst Appl. 2012;150(1):80–99 (In Chinese).
2. Zhang YX, Li S. Multiple neural network model based on data partition using feature clustering. Inf Control. 2013;42(6):693–9. (In Chinese).

3. Liu Y. A new algorithm of image segmentation based on bidirectional search pulse-coupled neural network. International Conference on Computational Aspects of Social Networks. IEEE; 2010. p. 458–63.
4. Liu Y. Image segmentation using artificial intelligence approaches. Electron World. 2013;119(3):38–41. (In Chinese).
5. Ghamisi P. An efficient method for segmentation of images based on fractional calculus and natural selection. Expert Syst Appl. 2010;39(1):12407–17.
6. Baßstürk A. Efficient edge detection in digital images using a cellular neural network optimized by differential evolution algorithm. Expert Syst Appl. 2009;36(8):2645–50.
7. Liu Y. A novel method for image segmentation based on nature inspired algorithm. Adv Intell Comput Technol Appl. 2010;4(1):7–13.
8. Guo R, Pandit SM. Automatic threshold selection based on histogram modes and discriminant criterion. Mach Vis Appl. 1998;10(5):331–8.
9. Sathya PD, Kayalvizhi R. Modified bacterial foraging algorithm based multilevel thresholding for image segmentation. J Eng Appl Artif Intell. 2011;24(4):346–52.

Chapter 73
Compiling Missile Environmental Spectrum and Accelerated Corrosion Test Method

Yuhao Shan, Jianli Shi and Wenhui Peng

Abstract Corrosion is an important factor in the missile reliability. The service environment of submarine-launched missiles may directly affect their life span. The corrosive environment study is of great significance to improve the reliability of the missile. Based on the specific characteristics of the local environment, it is proposed to accelerate the corrosion test on key parts with the help of environmental spectrum. Equivalent conversion method is used to establish the relationship between the corrosion equivalent acceleration and actual corrosion. KH-7700 three-dimensional digital video microscope is adopted to handle the corrosion surface of the objects.

Keywords Environmental spectrum · Corrosion · Equivalent conversion · Accelerated corrosion test

73.1 Introduction

Before the 1950s, the definition of corrosion was confined to metal corrosion. From the 1950s, many authoritative scholars or erosion research institutions tend to extend it to all materials, but it usually refers to the metal damage. Because the metal and its alloys are still the most important structural materials, the metal corrosion is one of the most interesting questions. The corrosion of submarine-launched missiles directly affects their reliability, and because of the complexity of its environment, a comprehensive study of the corrosion process of various factors is necessary. With the establishment of the spectrum under the actual environmental conditions of missiles and the equivalent conversion, we simulated the corrosion process of the specimens in laboratory. The local environment spectrum of missiles was compiled with reference to compilation of spectrum of aircraft-launched missiles [1]. Prior to the compiling of the physical environmental spectrum in which the elements of the study were analyzed, we selected the acquisition factor and time, which is necessary for the statistical data collection, together with equivalent conversion method

Y. Shan (✉) · J. Shi · W. Peng
Navy Submarine Academy, 266000 Qingdao, China
e-mail: juntong123@yeah.net

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_73

to accelerate the establishment of relations as the gist of the corrosion research in the laboratory. The research has great significance for the study of metal corrosion and protection methods.

73.2 Compiling of the Environmental Spectrum

The corrosive environmental factors of submarine-launched missiles can be divided into climatic elements and chemical environmental factors. There are some relationships between fatigue load spectrum and fatigue life of individual aircraft [2]. We establish the environmental spectrum to study the relationship between environment and corrosion. The main factors of climatic corrosion are temperature, humidity, salt spray, condensation, precipitation (moisture), and solid sediments, of which the temperature (T), humidity (RH), salt spray, and condensation are the most important corrosive factors. The chemical corrosion factors are mainly SO_2, SO_4^{2-}, nitrogen oxides NO_x, acid rain, salt water, salt spray, and halogen ions, in which the salt spray must be considered. The structural material of test object is the cast aluminum, which corresponds to the critical relative humidity of 65%.

Some researchers have already compiled the marine environmental spectrum [3]. According to the corrosion characteristics and environmental conditions of the object of study, we generally divide temperature into seven sections, 0–10 ºC, 10–15 ºC, 15–20 ºC, 20–25 ºC, 25–30 ºC, 30–35 ºC, and above 35 ºC, in the preparation of the spectrum environment. In terms of the division of temperature, the statistics of time and frequency of `precipitation, fog, and condensation correspond to different segments, as well as the action time when the relative humidity is more than 65% and the temperature exceeds 20%.

Researchers have already measured the air pollution density [4]. We imitated this thought. The first thing we did is to collect the abovementioned elements' data by using specific collection equipment. Then, through the preparation of environmental data processing procedures, the environmental data analysis and processing were carried out. Finally, the month spectrum and year spectrum were obtained.

Given the division of temperature, we accounted time, frequency, and concentration of precipitation, fog, and condensation in each month at the temperatures 5 ºC, 10 ºC, 15 ºC, 20 ºC, 25 ºC, 30 ºC, and 35 ºC. We also reported the temperature above 20 ºC and relative humidity more than 65% corresponding to the duration of action at all levels.

On the basis of the month spectrum, we assumed that certain environmental factors for each month are a random variable, so statistical value of the mean, standard deviation, and reliability that reaches 90% each month can be obtained over the years. Due to the special nature of the study, we selected 6 months in which the environment is relatively poor in a year to establish the environmental spectrum. These 6 months represent the actual environment of the study in this year. We accounted the statistics of environmental spectrum (Table 73.1) in a year through of data collection.

Table 73.1 Spectrum of rain, fog, condensation, and humidity

Temperature		20 °C	25 °C	30 °C	35 °C	Frequency
Spectrum of rain/h		124.5	114.4	45.9	1.9	79.4
Fog and condensation/h		205.2	85.0	2.8	0.6	59.3
Spectrum of humidity/h	70%	332.6	412.9	24.7	0	
	80%	687.6	821.5	29.0	4.1	
	90%	838.6	887.6	16.5	0	

73.3 Equivalent Conversion Method of Environmental Spectrum

Researchers have already shown the conversion principles and accelerated test [4–7]. Current I_C represents the corrosion rate of the metal. For the same metal member (the surface area is S), the electric current of corrosion from time t_1 to time t_2 can be expressed by the integral form: $Q = \frac{1}{F}\int_{t_1}^{t_2} I_c(t)\mathrm{d}t$.

According to the corrosion damage equivalence principle, for the same material, the electric quantity of corrosion in the actual conditions is equal to the electric quantity when under laboratory conditions, namely: $Q = Q'$.

Upon derivation, $t' = \alpha \cdot t$.

t' is the test time in the laboratory, t is the time in the environmental spectrum through collection, and α is the equivalent conversion factor.

The equivalent conversion factor can be expressed as: $\alpha = \frac{I_c}{I_{cp}}$

I_c is the electric current in any condition, I_{cp} is the electric current in standard condition.

When the standard condition changes, $I_{cp} \to I_{cp1}, \alpha_A \to \alpha_A'$, then: $\alpha_A = \frac{I_c}{I_{cp1}} \bullet \frac{I_{cp}}{I_{cp}} = \frac{\alpha_A}{\alpha_{cp1}}$

α_A is the equivalent conversion factor in standard condition and α_A' is the new equivalent conversion factor when the condition changes. Table 73.2 describes the corrosive equivalent conversion factor of missile aluminum alloy in the actual condition and in the laboratory condition. Tables 73.3 and 73.4 are the corresponding conversion factors of different concentrations of NaCl, HCl and H_2SO_4 with respect to water. Figure 73.1 is the equivalent conversion coefficient surface of aluminum alloy in water which is based on the calculated coefficients.

The specific conversion method:

Find out the corresponding equivalent conversion factor α_i of T_i, RH_i and α_p which corresponds to the standard condition T_p, RH_p Then

$$t_{pi} = \frac{t_i \bullet \alpha_i}{\alpha_p}, t_p = t_{p1} + t_{p2} + t_{p3} + \cdots + t_{pi} + \cdots + t_{pk} = \sum_{i=1}^{k} \frac{t_i \bullet \alpha_i}{\alpha_p} = \frac{1}{\alpha_p}\sum_{i=1}^{k} t_i \bullet \alpha_i,$$

where t_p is the total hours of accelerated corrosion test in the specified temperature and humidity.

Table 73.2 Corrosive equivalent conversion factor of aluminum alloy (α)

T (°C)	20	25	30	35	40
RH = 60%	0.10424	0.13591	0.19570	0.33044	0.66400
70%	0.16380	0.24440	0.32500	0.62526	0.92552
80%	0.16707	0.29053	0.42600	0.73177	1.03753
90%	0.11045	0.20700	0.29995	0.64997	1.00000

RH relative humidity

Table 73.3 Conversion factor of different concentrations of NaCl with respect to water (β)

Concentrations of NaCl(%)	0.35	0.5	3.5	7
	0.643	0.612	0.121	0.097

Table 73.4 Conversion factor of different concentrations of HCl, H_2SO_4, with respect to water (β)

Concentrations	1:10000	1:1000	1:500
	0.454	0.235	0.110
	0.635	0.348	0.302

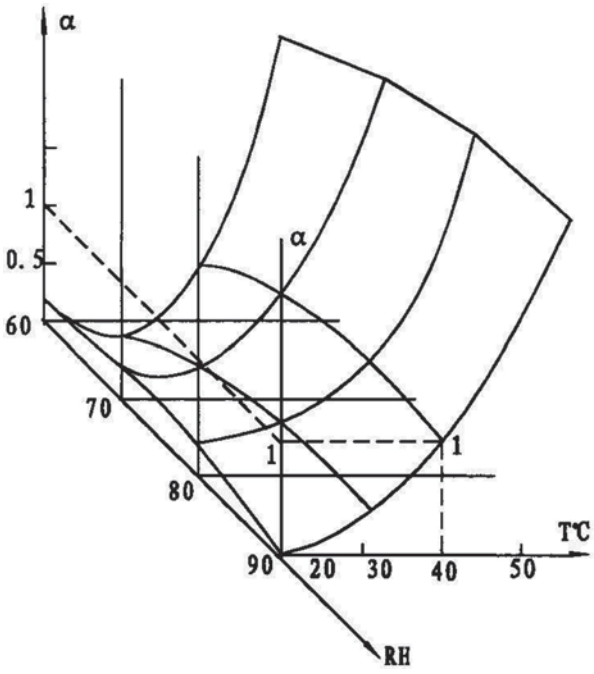


Fig. 73.1 Equivalent conversion coefficient surface of aluminum alloy in water condition

73.4 Environmental Spectrum of Test

The accelerated corrosion test was configured as: 3.5 % NaCl solution was added to a small amount of sulfuric acid to make pH = 4 ± 0.2, the temperature: T = 40 ºC, the humidity: RH = 90 % [8–10].

Based on Tables 73.1, 73.2, 73.3, and 73.4, we took the condition of T = 40 ºC and RH = 90 % as the benchmark of environmental conditions and calculated the environmental equivalent acceleration relations including the acting time of moist air of 659.05 h, rainfall time of 52.4 h, and reaction of salt spray and condensation of 41.45 h.

The conversion coefficient of concentration of 3.5 % NaCl solution is 0.121 with respect to moist air, and the acceleration coefficient is 8.26.

The concentration of sulfuric is 4 mg/L, as pH = 4. Through Table 73.4, the conversion coefficient of sulfuric acid is 0.2148, and its acceleration factor is 4.5783 when its concentration is 4 mg/L.

From the above, the total time of accelerated corrosion solution is

659.05/(4.5783 + 8.26) + 52.4/(4.5783 + 8.26) + 41.5/(4.5783 + 8.26) = 51.325(h).

Due to the special nature of the study, it is not necessary to consider the impact of rainfall, salt spray, and condensation. The total duration of action is

659.05/(4.5783 + 8.26) = 51.325(h).

We selected the calculated time as one period in accelerated corrosion test which can represent 1 year of the missile's actual environment.

73.5 Accelerated Corrosion Test

The test material is ZL115-T5 aluminum alloy at the geometric dimensions of 105 × 90 × 10 mm.

The test process includes the test pretreatment, the immersion test, the after treatment, the corrosion morphology and status determination, and the data processing. The accelerated corrosion test includes five cycles in total, and the configuration of corrosion solution is based on Chap. 4. Table 73.5 reveals the responding treatment of different specimens.

Table 73.5 Number and treatment of specimens

Number	Treatment
01, 09	No
02, 10	Smear corrosion X (MIL C 81309 E TYPE II)
03, 11	Smear soft film corrosion inhibito
04, 12	Smear epidural corrosion inhibito
05, 13	In touch with brass
06, 14	Smear corrosion X and scratch
07, 15	Smear sealant and scratch
08, 16	In touch with brass coated with lubricating oil

The specimens are vertically placed in a thermostatic glass tank and the cycle time is started. During the test, the test pieces for each period need the related measurement, including the measurement of the specimen weight by using the electronic balance and the specimen surface by using a digital three-dimensional microscope.

KH-7700 digital three-dimensional microscope was used to photograph the depth variation [7] of the same specimen in the five periods. The depth variation of 01 specimen in the first four periods is shown in Figs. 73.2 and 73.3. The depth of pit in the specimen surface is used to reflect corrosion. Figure 73.4 reveals the depth of pits on different specimens in five periods.

Analysis Employ the depth of pit in the specimen surface to reflect corrosion. In fig. 73.4, the depth of pit of the fourth specimen is basically unchanged and constant at zero. The depth of pit of the first, fifth, and eighth specimen is growing rapidly. As to the corrosive control of ZL115-T5 aluminum alloy, the TFHS-15 subdural corrosion inhibitor performs best and the effect of anticorrosion soft film inhibitor

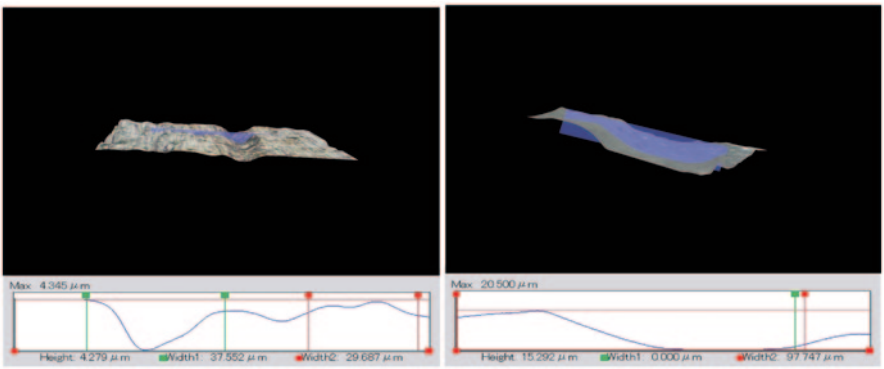


Fig. 73.2 Three-dimensional depth variation of 01 (*left:* the first period, *right:* the second period)

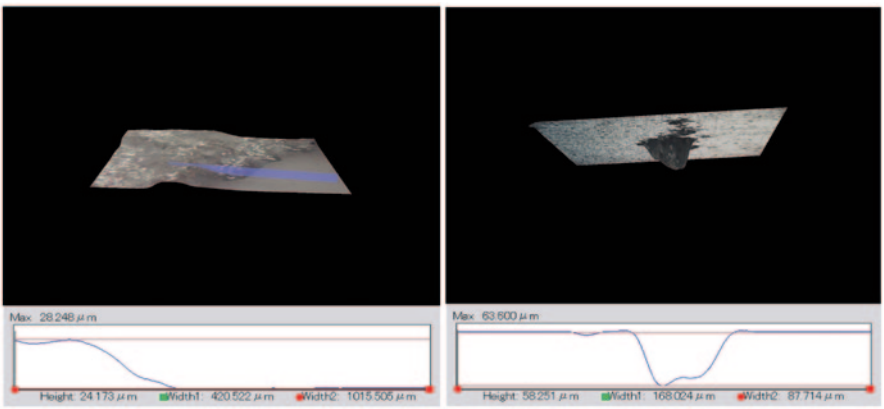


Fig. 73.3 Three-dimensional depth variation of 01 (*left:* the third period, *right:* the fourth period)

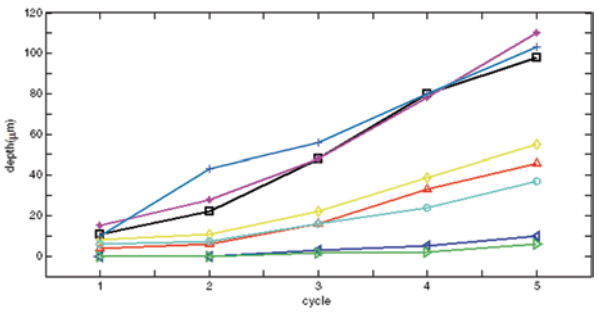


Fig. 73.4 Corrosion pit depth variation of specimens

TFHS-10 is second to TFHS-15 inhibitor. The effect of corrosion inhibitor, corrosion X is not obvious; therefore, TFHS-15 inhibitor and TFHS-10 inhibitor can be used as corrosion inhibitors for aluminum alloys to slow down or even suppress the emergence and development of missile corrosion.

73.6 Conclusion

For the first time, we study the corrosion process of submarine-launched missile by means of establishment of its environmental spectrum. On the basis of collection of the missile environmental data, the compiling spectrum is innovative. After the specimens were tested, we used the degree of pitting corrosion depth to reflect various degrees of corrosion of different processing specimens only to find that different inhibitors perform different roles. Besides, it is feasible to apply the corrosion control method; nevertheless, we have to further study the role of other corrosion protection measures in the mitigation of corrosion of materials, such as EXP-X, which combines the excellent thermal insulation and the corrosion resistance. These anticorrosion measures are expected to be applied to other areas.

References

1. Li Y, Wang Z, Chen YL et al. Research on compiling fatigue load spectrum of individual aircraft and analysis of fatigue life based on flight data[C]//Prognostics and System Health Management (PHM), 2012 IEEE Conference on. IEEE, 2012: 1–5.
2. Zhang TF, Zhao WY, Yang XH et al. Compiling the marine environment spectrum based on inversion satellite remote sensing data[C]//Intelligence and Security Informatics (ISI), 2011 IEEE International Conference on. IEEE, 2011: 354–7.
3. Li H, Cao Q, Song Z. Corrosion failure analysis about package box of aluminum silicon alloy used for microwave module[C]//Physical and Failure Analysis of Integrated Circuits (IPFA), 2013 20th IEEE International Symposium on the. IEEE, 2013: 578–82.

4. Zhang Y, Chang X, Zhang S et al. Probability model of surface corrosion damage ratio of aluminum alloy[C]//Education Technology and Computer, 2009. ICETC'09. International Conference on. IEEE, 2009: 212–4.
5. Yang XH, Jin P, Chen YL. Making out aerial environment spectrum of aircraft [J]. Acta Aeronautica et Astronautica Sinica. 2008;29(1):85–90.
6. Song EP, Liu WT, Yang X. Study on accelerated corrosion test environment spectrum for internal aircraft structure [J]. Acta Aeronautica et Astronautica Sinica. 2006;27(4):643–9.
7. Liu WT, Li YH, Chen QZ. Accelerated corrosion environmental spectrums for testing surface coatings of critical areas of flight aircraft structures [J]. J Beijing Univ Aeronaut Astronaut. 2002;28(1):109–12.
8. Zhou XY. Equity environmental spectrum and speed test spectrum for aircraft structure [J]. Acta Aeronautica et Astronautica Sinica. 1996;17(5):613–6.
9. Liu YH, Ren SY. Study on equivalent accelerated corrosion test environment spectrum of typical marine atmosphere [J]. Equip Environ Eng. 2011;8(1):48–52.
10. Chen YL, Duan CM. Local environmental spectrum for accelerated corrosion of aircraft structure [J]. J Nanjing Univ Aeronaut Astronaut. 1999;31(3):338–41.

Chapter 74
Aerosols' Statistical Fractal Characteristics Based on Their Shape Parameter in Optical Sensors

Juan Yang and Fang Gu

Abstract Based on the shape structure and the pulse signal's characteristic parameters, aerosols' statistical models are established in this chapter. The counting distributions of aerosols' scattering pulse signals are measured by using a counting device with 2048 uniform voltage channels. The experimental results show that the counting distributions obey the lognormal distribution with the natural number as the independent variable. On this basis, the relations of counting distributions between the aerosols' projection area characteristic parameters and the scattering pulse signals' characteristic parameters are discussed. Then, the statistical fractal characteristic of the pulse signals' different parameters is deduced; besides, the physical meaning and the calculation method of the fractal dimension are presented.

Keywords Fractal · Shape parameter · Pulse signal · Lognormal distribution · Aerosol

74.1 Introduction

The fractal theory proposed by Mandelbrot in 1975 [1] has been widely used in the research of all kinds of science [2, 3], and its common mathematical representation is the power function distribution [1]. Many real signals have obvious fractal characteristics. Namely, there is a natural connection between the fractal and the signal. This connection establishes the foundation of fractal theory used in the field of the signal processing [3].

As to aerosols' scattering pulse signals, these formations must be received by a series of random factors. Considering the pulse signal amplitude distributions of different pulse widths, these distributions must obey the same statistical distribution laws because its random factors in the formation process are the same; that is,

J. Yang (✉)
School of Information and Technology, Jinling Institute of Technology, 211169 Nanjing, China
e-mail: yangjuan@jit.edu.cn

F. Gu
School of Physics and Optoelectronic Engineering, Nanjing University of Information Science and Technology, 210044 Nanjing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_74

there is a certain similarity. In this sense, there must be some connection between the counting method pulse signal distribution and the fractal; but now there is no theoretical basis for the fractal structure between the characteristic parameters of different counting pulse signals.

On this basis, aerosols' statistical models of their shape and pulse signal parameters are established first in this chapter. Then, we use a 2048-uniform voltage channel counting device to measure the particles' scattering pulse signals, with the result showing that the counting distribution of the pulse signal group matches well with the logarithmic normal distribution with a natural number as the independent variable with high accuracy. Besides, the relation of the distributions between the projected area characteristic parameters and the scattering pulse signal characteristic parameters is discussed. In this sense, we can deduce that there is a statistical fractal characteristic among different parameters of the counting pulse signals. Finally, the physical meaning and the calculation method of the fractal dimension are proposed.

The rest of this chapter is organized as follows: First, the characteristic parameters and their statistical distributions of the aerosol and the random signal for aerosol are introduced; then, the characteristic parameters' statistical distribution of the random signal is given; finally, the statistical fractal characteristics of these distributions and their fractal dimension are discussed.

74.2 Characteristic Parameter and Its Statistical Distribution of the Aerosol

The particles that occur naturally usually have different shapes. In order to characterize the particle size, there are several kinds of measurement principles related to the marking method of particle size [4, 5], such as sieving diameter, volume diameter, and diameter of the surface area. As the particles usually feature irregular shapes, the particles on the projected area in different directions are not the same. Herein, we take the maximum projected area of the particles as the basic parameters of the particles in this chapter.

Mass is one of the basic characteristics for suspended particulates. Any particle group N can be classified according to the size of its mass, based on which the particles can be divided into "mass subset" $N(m)$ with different levels. The particles can be generated by a random process with the same quality, but the geometry can vary widely; therefore, the projected area of particles is very wide.

The particles are typically a random group if we sort the particles according to the size of their projected areas. Let the minimum projected area of the mass subset be S_m, and the maximum projected area of the mass subset be S_M, then any particles' projected area of the mass subset satisfies $S_m \le S \le S_M$.

Considering the actual measurement accuracy, we can also use a natural number l_S to describe the mass subset. As to the jth particle of the mass subset m, there is the following equation:

$$l_S = \left\lceil \frac{S_j - (S_m - \Delta S)}{\Delta S} \right\rceil \tag{74.1}$$

Obviously, based on the natural number l_S, the pulse counting distribution functions of mass subset can be expressed as $f(l_S) \equiv \lim_{N\to\infty} \frac{n_{l_S}}{N}$, where N represents the total number of particles, and n_{l_S} denotes the number that the projected area of the particle is l_S. Considering the knowledge of the statistics, the distribution functions $f(l_S)$ will tend to be steady when the sample's number of subsets is large enough in the actual measurement.

74.3 Characteristic Parameter and Its Statistical Distribution of the Random Signal for Aerosol

In this chapter, we use the laser airborne particle counter, whose sampling rate is 28.3 L/min to measure the statistical distribution of the aerosol's scattering pulse signals. The whole measurement system can be seen in Fig. 74.1.

It is obvious that the output signal of photo-electricity in airborne particle counter is a time series $f(t)$, and a subsample of $f(t)$ can be seen in Fig. 74.2 [6]. When the sample intervals of the signals are certain, the continuous signal $f(t)$ can be written as $f(i{\cdot}\Delta t)$, which is a natural number sequence. Taking a reference value $V_0(f_{Max} > V_0 > f_{\min})$, a two-parameter transformation model, whose transformation function is $\tilde{f}(t, V_0) = f(t) - V_0 > 0$, can be defined to characterize the statistical characteristics of $\{f(i{\cdot}\Delta\tau)\}$.

Obviously, the set $\{\tilde{f}(t, V_0)\}$ is composed of pulse sequences, which have discontinuous time sequence. On this basis, a random pulse $\tilde{f}_j(t, V_0)$ composes a pulse structure whose amplitude is defined as V_j, and its width is defined as τ_j, which is shown in Fig. 74.2.

If we use $\Delta\tau$ and ΔV to represent the actual accuracy of the measurement devices, then a set of natural numbers (l_τ, l_V) can be used to describe the random pulse signal [6], which is expressed as follows:

$$l_\tau = \left\lceil \frac{\tau_j}{\Delta\tau} - 1 \right\rceil \quad l_V = \left\lceil \frac{V_j}{\Delta V} \right\rceil \tag{74.2}$$

According to Eq. (74.2), the probability distribution function that the measurement results appear in signal amplitude channel can be expressed as:

$$q(l_V) \equiv \lim_{N\to\infty} \frac{n_{l_V}}{N} \tag{74.3}$$

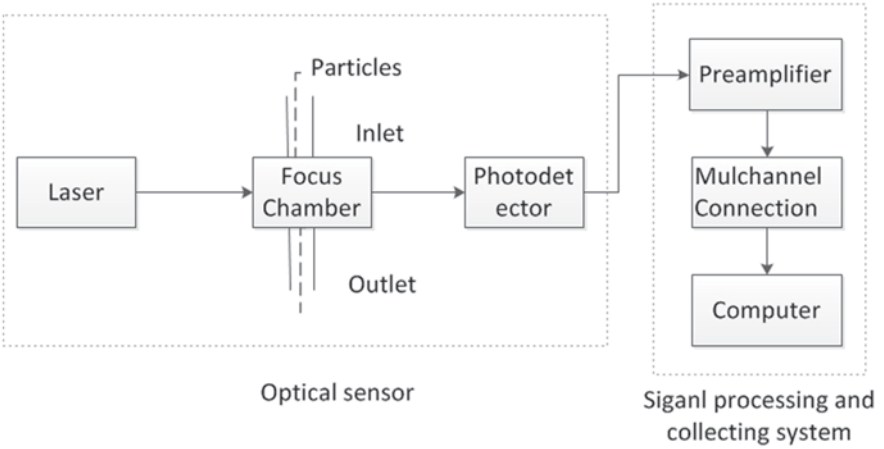
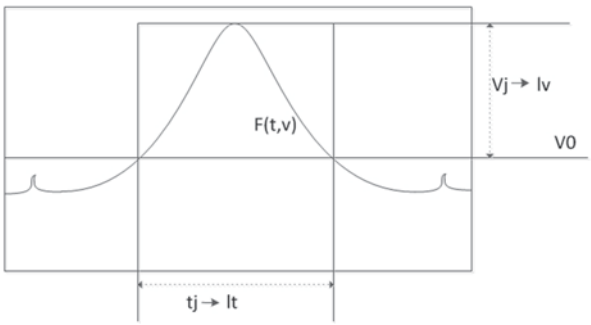


Fig. 74.1 The whole measurement system of laser airborne particle counter

Fig. 74.2 Sketch map of the pulse structure



where N denotes the total number, and n_{l_V} denotes the pulse numbers that the measurement results appear in the $l_V th$ signal amplitude channel. When the sample number is large enough, the distribution functions $q(l_V)$ will tend to be steady. Similarly, the measurement results can also be described by the corresponding parameters τ_M, τ_m and the probability distribution function $p(l_\tau) \equiv \lim_{N\to\infty} \frac{n_{l_\tau}}{N}$, where n_{l_τ} denotes the pulse numbers that the measurement results appear in the $l_\tau th$ signal width channel.

74.4 Characteristic Parameters' Statistical Distribution of the Random Signal

The statistical distribution functions of aerosols' pulse signals $q(l_V)$ and $p(l_\tau)$ are the projection that measure results are l_V and l_τ in the same environment; therefore, they must have the same nature. In this sense, if the nature of the random effect is not changed, the type of the one-dimensional statistical distribution function related to this nature will be unchanged. On this basis, we can conclude that one-dimensional distribution functions $q(l_V)$ and $p(l_\tau)$ have the similarity because they reflect the common characteristics of the random effect. Based on this logic, we can also conclude that the counting distribution of particle projected area $f(l_S)$ has similarities with the functions $q(l_V)$ and $p(l_\tau)$ because of the statistical randomness under the same environmental system.

With the above device, taking the serial width scope $\tau_k \sim \tau_k + \Delta\tau$, the measured counting distribution function $q(l_V)$ of the random pulse signal amplitude can be shown as discrete points in Fig. 74.3 [6].

As shown above, the function $q(l_V)$ has an obvious center asymmetry; besides, the nonlinearly similar characteristic between different function sequences is seen quite well. Considering the signal amplitude distribution orderliness generated by the standard particle group and the statistical distribution law of particle size

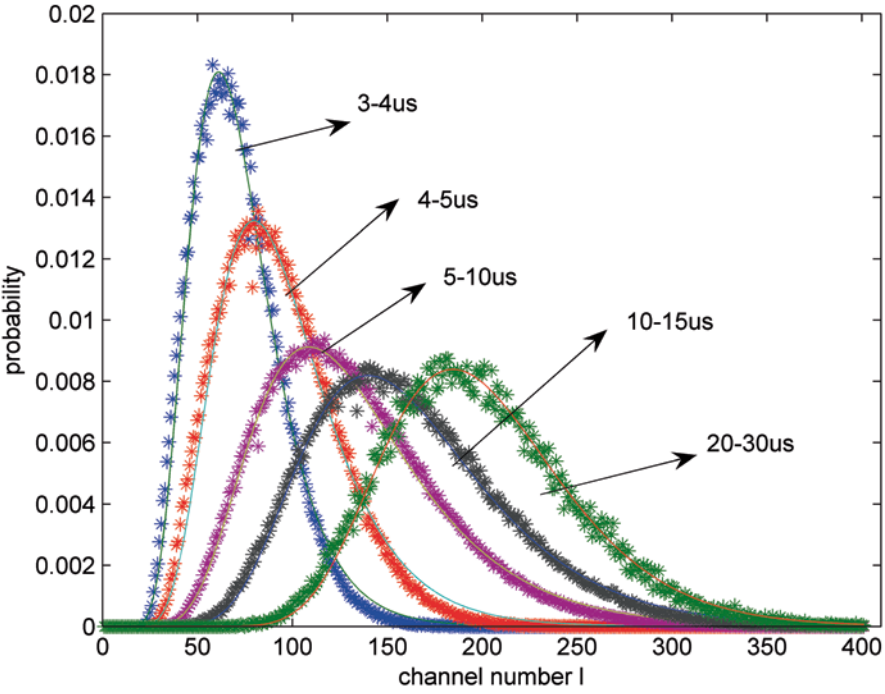


Fig. 74.3 Scattering signal amplitude distributions under different signal widths

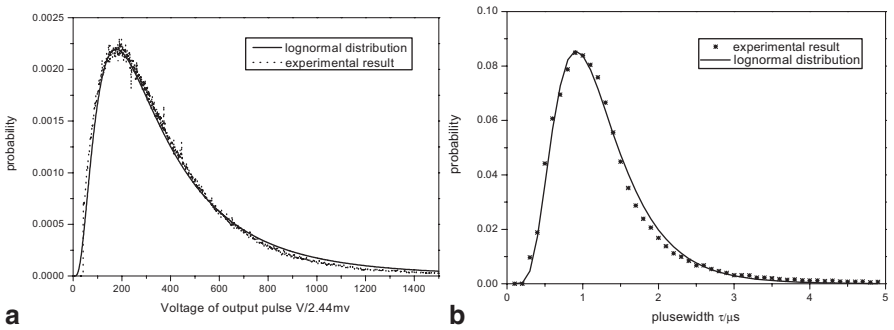


Fig. 74.4 Amplitude and width distributions of pulse signal

obtained by high-precision measurements, we choose a statistical function with the form of the lognormal distribution to fit the measured data with the fitting results, seen as a sequence of solid lines in Fig. 74.3. The equation corresponding to the solid line in Fig. 74.3 satisfies [6]:

$$q\left(l_V\right) \equiv \frac{1}{\Omega \sigma_{\ln} l_V} e^{-\frac{\left(\ln l_V - \mu_{\ln}\right)^2}{2\sigma_{\ln}^2}} \tag{74.4}$$

where $\mu_{\ln} = \overline{\ln m}$, $\sigma_{\ln} = \sqrt{\overline{(\ln m - \overline{\ln m})^2}}$, and Ω is the normalization coefficient [7]. Formally, Eq. (74.4) corresponds to the lognormal distribution function $q(l_V)$, but the independent variable is the natural number and the normalized coefficient is not $\sqrt{2\pi}$.

The width and amplitude distributions of the random pulse signal group for the particle scattering are measured by the same device, and the results appear in Fig. 74.4. The total channel number of amplitude distribution is almost $L \approx 1500$, and the statistical results can be given directly in Fig. 74.4a. Then, the total channel number of width is almost $K \approx 50$, while its distribution uses the subsection measurement form. Maintaining the invariability of the total sample and successively increasing the lower limit of the pulse width, the counting rate in the channel of different widths is calculated, as shown in Fig. 74.4b [8]. As seen from Fig. 74.4, the distribution curves of the amplitude and width have geometric similarity.

The calculation formula of fitting curves can be expressed as follows:

$$p(l_\tau) \equiv \frac{V_M - V_0}{\Omega_V \Delta V} \frac{1}{\sigma_{\ln} l_\tau} e^{-\frac{\left(\ln l_\tau - \mu_{\ln}\right)^2}{2\sigma_{\ln}^2}} \tag{74.5}$$

$$q(l_V) \equiv \frac{\tau_M}{\Omega_\tau \Delta \tau} \frac{1}{\sigma_{\ln} l_V} e^{-\frac{\left(\ln l_V - \mu_{\ln}\right)^2}{2\sigma_{\ln}^2}} \tag{74.6}$$

where the fitting results can be seen as the sequence of solid lines in Fig. 74.4.

According to the above analysis, we can see that the signal amplitude and the width counting probability distributions all obey the lognormal distribution function with the natural number as the independent variable when we use a two-parameter analyzing model to deal with the measurement results of the random signal generated by the aerosol.

As the counting distribution of the particle projected area $f(l_S)$ has similarities with functions $q(l_V)$ and $p(l_\tau)$, and they have good nonlinear scaling invariance between functions $q(l_V)$ and $p(l_\tau)$, it can be concluded that they have good nonlinear scaling invariance between the counting distribution of particle projected area $f(l_S)$ and the signal counting distribution functions $q(l_V)$ and $p(l_\tau)$; that is to say, the counting distribution of particle projected area $f(l_S)$ also obeys the lognormal distribution function with the natural number as the independent variable.

74.5 Statistical Fractal Characteristics of the Distribution Function and Its Fractal Dimension

Thinking from the logical form, the normalized coefficient of the probability distribution function will tend to be constant if $\Delta V, \Delta\tau \to 0$ and $L_V, L_\tau \to \infty$. According to Eq. (74.4), the signal counting distribution curves have nonlinear scaling invariance. On this basis, if we perform the nonlinear transformation $l_V = (bl_\tau)^\beta$ between the two curves, there must be parameters b and β which can make the two transformed curves coincide. It can be described as follows:

$$q(l_V)dl_V = \frac{dl_\tau}{\sqrt{2\pi}\left(\frac{\sigma_{\ln V}}{\alpha}\right)l_\tau} e^{-\frac{(\ln l_\tau + \ln b - \mu_{\ln V}/\alpha)^2}{2\sigma_{\ln V}^2/\alpha^2}} = p(l_\tau)dl_\tau = \frac{dl_\tau}{\sqrt{2\pi}\sigma_{\ln\tau} l_\tau} e^{-\frac{(\ln l_\tau - \mu_{\ln\tau})^2}{2\sigma_{\ln\tau}^2}} \tag{74.7}$$

where $\beta = \sigma_{\ln V}/\sigma_{\ln\tau}$ and $b = e^{\mu_{\ln V}/\beta - \mu_{\ln\tau}}$. It can also be expressed by the corresponding physical parameters as follows:

$$\frac{V}{\Delta V} = \left[\left(\frac{\tau}{\Delta\tau} - 1\right)b\right]^\beta = \left[\left(\frac{\tau}{\Delta\tau} - 1\right)e^{\frac{\sigma_{\ln\tau}}{\sigma_{\ln V}}\mu_{\ln V} - \mu_{\ln\tau}}\right]^{\frac{\sigma_{\ln V}}{\sigma_{\ln\tau}}} \tag{74.8}$$

Commonly, $\beta = \sigma_{\ln V}/\sigma_{\ln\tau}$ is not a natural number as $\sigma_{\ln V} \neq \sigma_{\ln\tau}$. Eq. (74.8) is also the nonintegral dimension fractal power function form of the statistical structure for the random pulse signal group, where the fractal dimension corresponds to the ratio of the lognormal dispersion $\sigma_{\ln}$ for the characteristic parameters' counting distribution [9].

Furthermore, we can replace s in Eq. (74.8) with V and τ as the projection area parameter S, then the power function relationship can be concluded among the parameters S, V, and τ. That is to say, the characteristic parameters' counting distribution functions $f(l_S)$, $q(l_V)$, and $p(l_\tau)$ highlight obvious statistical fractal characteristics.

74.6 Conclusion

In this chapter, the statistical distribution model of the aerosol based on its shape parameter and its pulse signal characteristic parameters is established. Then, the relations between the projection area parameter and the pulse signal parameter are also discussed. On this basis, the statistical fractal characteristics among the characteristic parameters' counting distribution functions $f(l_S)$, $q(l_V)$, and $p(l_\tau)$ are deduced; in addition, the physical meaning and the calculation method of the fractal dimension are also given.

Acknowledgements This work was funded by the National Natural Science Foundation of China (Grant No. 61302167 and No.61307113) and the Scientific Research Startup Foundation of Jinling Institute of Technology (Grant No: jit-b-201111).

References

1. Mandelbrot BB. How long is the coast of Britain. Science. 1967;156(3775):636–8.
2. Ren X, Guo L. Fractal characteristics investigation on electromagnetic scattering from 2-D Weierstrass fractal dielectric rough surface. Chin Phys B. 2009;17(8):2956–62.
3. Zhao J, Lei L, Pu X. Fractal theory and its application in signal processing. Beijing: Tsinghua University Press; 2008. p. 3–4 (in Chinese).
4. Allen T. Particle size measurement. Beijing: China Architecture & Build Press; 1984. p. 108–12.
5. Schmidt-Ott A, Wüstenberg J. Equivalent diameters of non-spherical particles. J Aerosol Sci. 1995;26(Suppl 1):S923–4.
6. Yang J, Bian B, Yan Z, et al. The model of two parameters statistical fractal for aerosol scattering pulse signal. Acta Phys Sin. 2011;60(1):010508-1–010508-6 (in Chinese).
7. Yang J, Bian B, Yan Z, et al. The statistical fractal characteristic of the aerosol scattering pulse signal group. Acta Phys Sin. 2010;59(11):7713–8 (in Chinese).
8. Peng G, Lai X, Yan Z, et al. The fractal characteristic study of the statistic parameter of particles' scattering pulse signal. Acta Opt Sin. 2010;30(6):1693–6 (in Chinese).
9. Yan Z, Bian B, Wang S, et al. Analysis of random laser scattering pulse signals with lognormal distribution. Chin Phys B. 2013;22(6):246–56.

Chapter 75
Image Retrieval of Poisonous Mushrooms Based on Relevance Feedback and Clustering Algorithm

Hongbo Luo and Sujuan Zhou

Abstract Poisonous mushrooms have no significant difference with the edible fungus. In this paper, image retrieval technology based on color, shape and texture characteristic are applied to identify poisonous mushrooms With relevance feedback mechanism and clustering algorithm used to improve the retrieval efficiency. Clustering can divide similar images into a same class and greatly reduce the range of image retrieval. From the retrieved image, representative clustering images are selected to the users who can choose a satisfactory clustering from them and provide feedback to the system. Experimental results with this method show effectiveness of identifying unknown poisonous mushrooms.

Keywords Poisonous mushroom · Image retrieval · Relevance feedback · Clustering algorithm

75.1 Introduction

Poisonous mushrooms are growing widely in China. According to incomplete statistics, more than 250 kinds of poisonous mushrooms have been found in China, among which, 112 kinds are found in Guangdong Province [1]. In the vast rural areas and mountain towns, cases of eating poisonous mushroom occur now and then. As some poisonous mushrooms have no significant difference with the edible fungus, the image retrieval technology is introduced to help recognition innovatively. Such features as color [2], texture [3] and shape [4] are often used in traditional image retrieval. Considering the characteristic of mushrooms, the combined features of Content Based Image Retrieval (CBIR) method is used in this paper. Combined features can achieve better results than single feature.

However, the features extracted automatically by computer cannot well represent its semantic content [5]. The human's participation can make up for the disad-

S. Zhou (✉) · H. Luo
College of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou Higher Education Mega Center, 510006, Guangzhou China
e-mail: susona2002@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_75

vantage; thus relevance feedback technique [6] was introduced into image retrieval. More images relevant to the query image can be found efficiently by the interactive learning and retrieval process. Based on all these technology, an assisted identification system of poisonous mushrooms has been built to help more people distinguish poisonous mushrooms and help find new species of mushrooms.

75.2 Image Feature Extraction of Mushroom

Currently, the image feature extraction is still based on low-level features [7], like color feature, texture feature, shape feature. The next will focus on the combined features extraction of poisonous mushrooms.

75.2.1 Colour, Texture and Shape Feature Extraction

As RGB (Red, Green and Blue) space is not so efficient for search, it needs to be converted into the corresponding HSV (Hue, Saturation and Value) values, as in Eq. (75.1). For the sake of less computational cost, H, S, V color channels are separately quantified. The color histogram features of all sub block are extracted and normalized [8].

$$
\begin{aligned}
h &= \begin{cases} \arccos \dfrac{(r-g)+(r-b)}{2\sqrt{(r-g)^2+(r-b)(g-b)}}, b \le g \\ 2\pi - \arccos \dfrac{(r-g)+(r-b)}{2\sqrt{(r-g)^2+(r-b)(g-b)}}, b \ge g \end{cases} \\
s &= \frac{\max(r,g,b)-\min(r,g,b)}{\max(r,g,b)} \\
v &= \frac{\max(r,g,b)}{255}
\end{aligned} \tag{75.1}
$$

Gray level co-occurrence matrix (GLCM) is used for the texture feature extraction. Images are quantized from 256 levels down to 16; then calculate the corresponding texture feature vectors, and normalize the feature matrix. The texture feature parameters based on GLCM include energy, contrast, entropy, and the correlation. Their definitions can be found in the references as quoted.

After calculating the four textures feature parameters, average u and variance σ are calculated as to each feature parameters, which can suppress the direction vector so that the resulting texture features cannot be affected by the direction. An 8-dimensional feature vector like Eq. (75.2) is generated to represent the image texture features.

$$f = (\mu_{ASM}, \mu_{CON}, \mu_{ENT}, \mu_{COR}, \sigma_{ASM}, \sigma_{CON}, \sigma_{ENT}, \sigma_{COR}) \tag{75.2}$$

The most commonly used method of describing the shape features include: moments described method [9], boundary description method and geometric parameter method.

Among all the methods, the moment is an important method, which describes the average gray level distribution of region. Hu proposed the theory of invariant moments in 1962. With a set of geometric moment's linearly independent group consisting of a non-linear moments, export the moment with a group of rotation, translation and scale invariance, known as Hu invariant moments [10].

75.2.2 Features Combination Method

As the image retrieval result based on single feature is not so ideal, color, texture, and shape feature combination method is used in order to get higher retrieval rate.

Combined features of similarity calculate like Eq. (75.3).

$$Dis(i, j) = W_t D(i, j) + W_c E(i, j) + W_s I(i, j) \tag{75.3}$$

i, j represents two images, Wt represents the weight of texture feature value, Wc represents the weight of the color feature value, Ws represents the weight of the shape feature value, and Wt + Wc + Ws = 1, D (i, j) represents the similar distance of two image texture features, E (i, j) represents the similar distance of color features, I (i, j) represents the similar distance of shape features.

The system uses a weighted combination of non-uniform features. Set Wt = 0.2, Wc = 0.7, Ws = 0.1.Experiments show that using the proportion of this factor can get better search results.

75.2.3 Similarity Matching

Determining whether two images are similar is based on the image features similarity matching. The extracted image features are expressed in the form of vectors; then calculate the distance between the feature vectors. The smaller the distance is, the greater similarity will be. Thus the similarity measure is the key step of image retrieval, which affects the search result.

There are a lot of methods of similarity measurement; but most of them just calculate distance of superposition between different features. In this paper, Euclidean distance is used. As in Eq. (75.4), x, y represents the two images feature vectors, xi, yi represents the feature component, D(x, y) represents the distance of the images [11].

$$D(x, y) = \left(\sum_{i=1}^{n} | x_i - y_i |^2\right)^{1/2} \tag{75.4}$$

75.3 Use of Relevance Feedback and Clustering

Clustering is a meaningful process which divides the data into multiple classes. It's an unsupervised automatic classification. Similar images will be divided into the same class, which may greatly reduce the range of image retrieval. The clustering method can also be used in the feedback process. From the retrieved image, representative clustering images are selected to the users who can choose a satisfactory clustering from them and provide feedback to the system.

75.3.1 K-means of Error Squared and Criterion Function

The steps of K-means algorithm based on error squared and criterion function are shown as below.

a. Given data sets containing n vectors point, set I = 1, select K initial cluster centers Zj (I), j = 1,2,…, K;
b. Calculated distance of each data point from the cluster center D (Xi, Zj (I)), i = 1,2,…, n, j = 1,2,…,K, if D (Xi, Zk (I)) = min {D (Xi, Zj (I))}, then Xi ∈ wk;.
c. Calculate the new K cluster centers $Zj(I+1)=\frac{1}{n}\sum_{i=1}^{n_j} X_i^{(j)}$, j = 1,2,…, K, error squared and criterion function Jc value, $Jc(I+1)=\sum_{j=1}^{K}\sum_{k=1}^{n_j}\left\|X_k^{(j)}-Z_j(I+1)\right\|^2$;
d. If | Jc (I + 1)-Jc (I) | < ξ, the algorithm ends, otherwise I = I + 1, return to (b)

With the adjustment of classification vectors point, the squared error criterion function Jc is constantly decreasing. When the value of Jc does not change or changes within the acceptable range, the classification will be optimal. By this algorithm, people can select the K clusters according to their demand.

75.3.2 Relevance Feedback

Relevance feedback is an instructive learning technique in information retrieval. Its goal is to learn from the user interaction with the query system to discover and capture the intent of the user's query so that the defects of query system can be fixed. The relevance feedback system can be modified in the query strategy so as to increase the image retrieval system's adaptive function.

Relevance feedback mechanisms are used in this paper with the application of clustering algorithm. Firstly, the system uses wavelet feature [12] to classify all images by K-means clustering. The database has 100 images, which are divided into 5 categories, 20 of each category. When users are searching images, the feedback (similar or dissimilar) would classify similar images directly to a class and the sample image without similar images are marked as uncategorized pictures, as shown in Fig. 75.1.

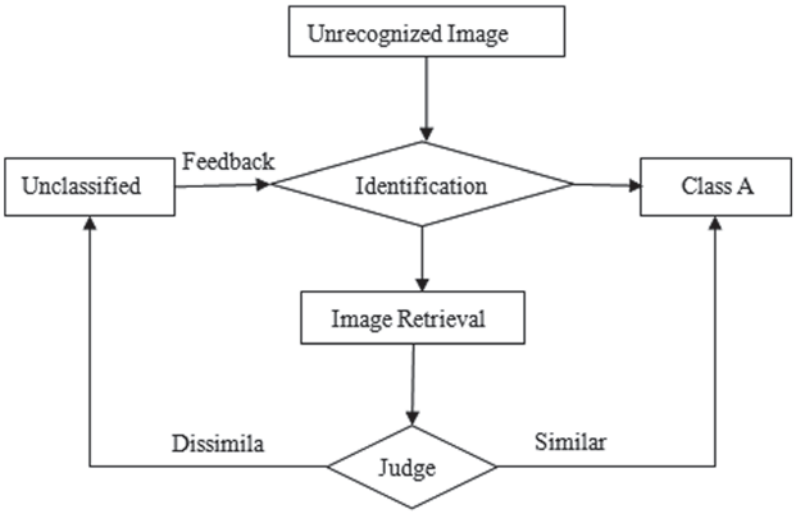


Fig. 75.1 Relevance feedback mechanisms

With the introduction of relevance feedback, it can easily cluster all images in the database and good search results will be obtained upon repeated searching and training.

75.4 Experiment Results and Analysis

Figure 75.2 shows the basic framework of poisonous mushrooms assisted recognition system based on relevance feedback and clustering algorithms. Firstly, collect images, and create image database; secondly, analyze the image in the database and extract the feature, create the image feature database; finally, analyze the sample images and get their feature vector. By similarity measurement, the query results return to the user according to the sorted similarity.

In the platform, all images are clustered into five major categories by K-means clustering. If the similarity is greater than 60 %, it can be determined that it belongs to the class in the category range. Experiments show the retrieval rate can be increased above 2 s with the K-means clustering (Fig. 75.4).

Seen from above, solely relying on the image features to do image retrieval cannot reach very good results. Users' experience of identifying poisonous mushrooms should be taken into account; therefore, it is necessary to introduce relevance feedback mechanism.

Upon studying and training for several times, Fig. 75.5 gets a satisfying retrieval result based on the combined features.

Compared with Fig. 75.3 and Fig. 75.5, we can see that the result is more desirable by adding relevance feedback. Upon repeated experiments, the image database can be used to identify the input images.

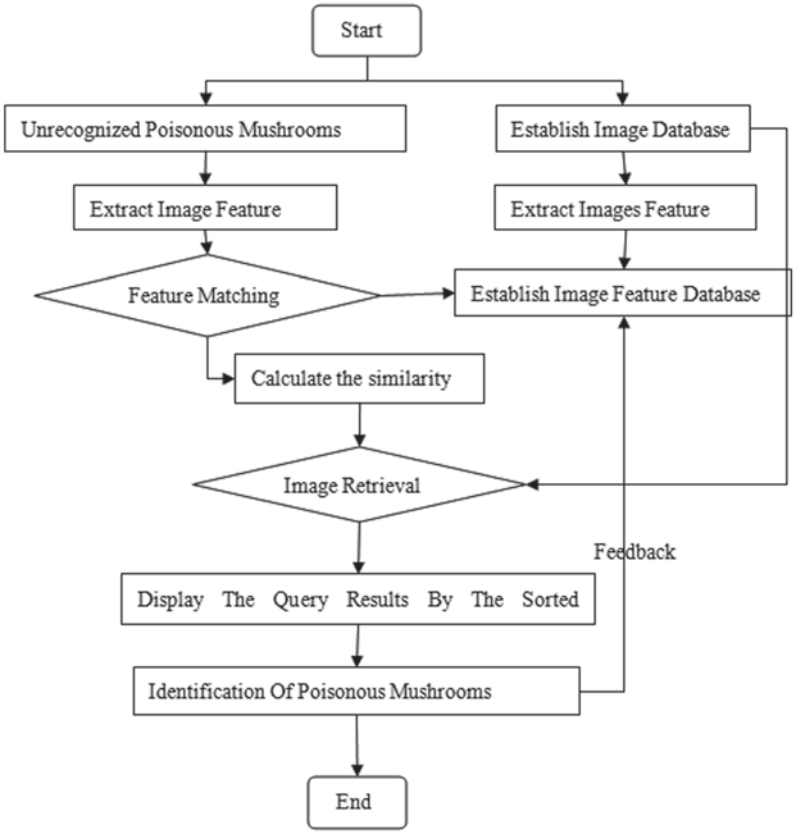


Fig. 75.2 Framework of system

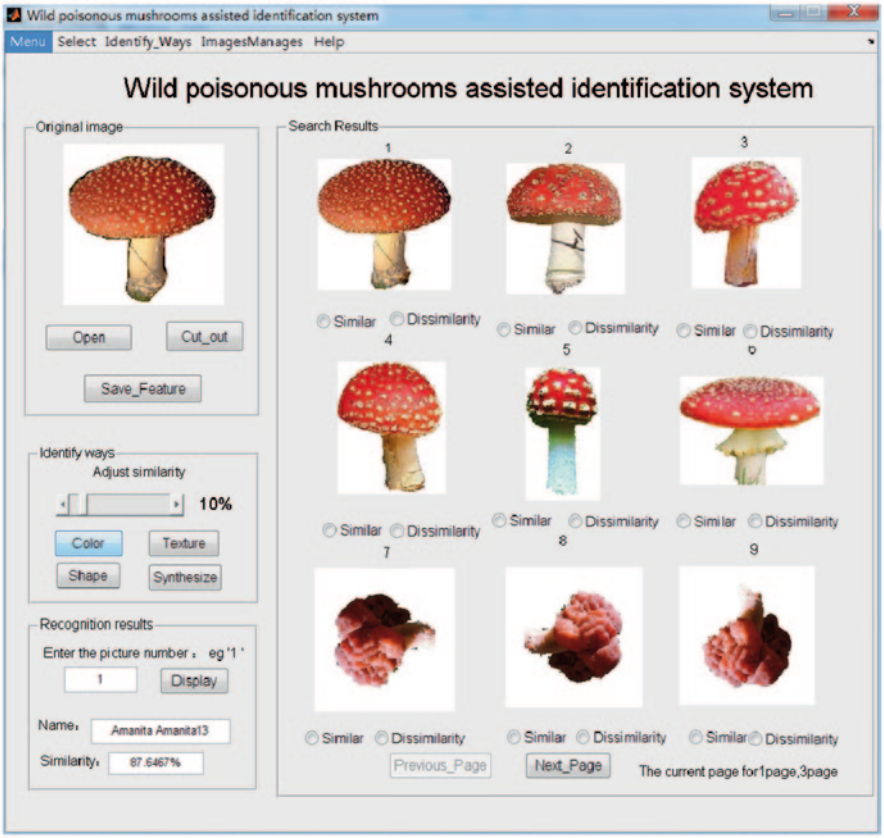


Fig. 75.3 Retrieval results based on color feature

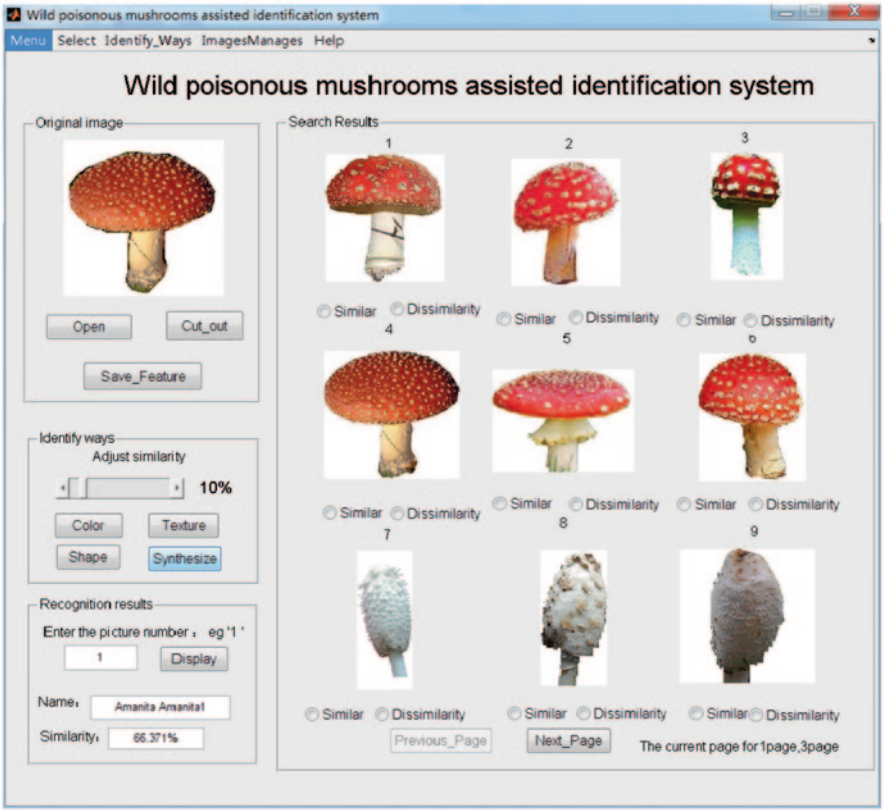


Fig. 75.4 Retrieval results based on combined features

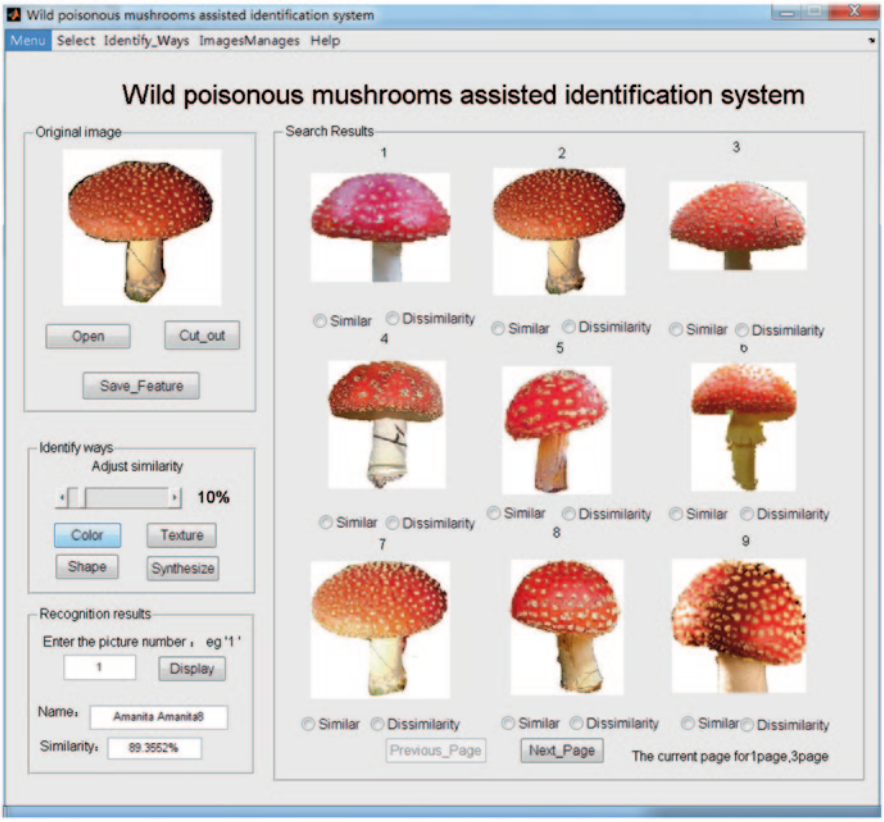


Fig. 75.5 Retrieval results based on color feature and relevance feedback

75.5 Conclusion

This paper studies on the identification of poisonous mushrooms based on image features, and discusses the key technology of content-based image retrieval, including the feature extraction methods and the similarity measurement. This paper also carries out research on the image feature extraction by means of relevance feedback mechanism and clustering algorithm. Experiments show these may help improve the retrieval efficiency. Above all, this paper has offered beneficial attempts for the identification of unknown poisonous mushrooms.

Acknowledgements This work is supported by the Innovation and Entrepreneurship Training Project for Undergraduate of Guangdong Province (1057313024)

References

1. Deng W, Li T. Guangdong known poisonous mushroom species. Fungal Res. 2005;4(1):7–12. (In Chinese).
2. Jiang L, Shen G, Zhang X. An image retrieval algorithm based on HSV color segment histograms. Mech Electr Eng Mag. 2009;26(11):54–7. (In Chinese).
3. Wang B, Yao H, Li B. An efficient image retrieval approach based on generalized image co-occurrence matrix. Wuhan Univ J Nat Sci. 2006;31(9):761–3. (In Chinese).
4. Zeng J, Wang J. Multi-feature image retrieval method based on K-means clustering segmentation. Comput Eng Appl. 2013;l49(2):227–30. (In Chinese).
5. Rui Y, Huang TS, Michael O, Mehrotra S. Relevance feedback: a power tool for interactive content-based image retrieval. [C] Circuits Syst Video Technol; IEEE. 1998;8(5):644–55.
6. Rui Y, Huang TS. A novel relevance feedback technique in image retrieval. [C] Proc 7th ACM Int Conf Multimed; 1999;2(3):67–70.
7. Petrou M. Image processing [M]. Beijing: Machine Press; 2005. p. 112–20.
8. Li J. A new image retrieval method based on color and texture features. Inf Technol Mag. 2007;5(3):84–6. (In Chinese).
9. Zhang J, Fu S. Shape recognition based on fourier descriptor of moments. Microcomput Inf. 2009;25(34):132–5. (In Chinese).
10. Yang L, Zhang F. An improved image retrieval algorithm. J Commun Univ China Sci Technol. 2011;4(3):25–8. (In Chinese).
11. Huang D, Hu J, Zheng Y. Development research of relevance feedback in interactive image retrieval. Appl Res Comput. 2005;9(2):15–8. (In Chinese).
12. Song H, LI B, Mu W. Retrieval algorithm based on wavelet image feature points. Comput Eng Des. 2007;28(2):386–9. (In Chinese).

Chapter 76
Development of an Elevator Energy Consumption Monitoring Device

Chuanglu Lin, Xinhua Wang, Xingjun Wu and Zhongxing Li

Abstract The energy efficiency of elevators has received more and more attention. There are few elevator energy consumption-monitoring devices for evaluating the energy consumption and generating energy-saving measures. To close this gap, this chapter presents the development of an elevator energy consumption-monitoring device, which uses the current transformer and voltage divider networks to sense the electrical signal. The second-order Σ−Δ analog-to-digital converters (ADCs) and specific digital signal processor (DSP)-based electrical measurement integrated circuit (IC) is selected to acquire electric parameters. The magnetic-coupled isolators are used to obtain fast digital communication and common mode rejection. Advanced RISC machines (ARM) processing architecture-based electric parameters acquisition, debugging console, power calibration, fast nonvolatile storage, and remote communication are designed. The experiment shows that the accuracy of the device is better than 0.2 %, which meets the design requirements; and it works stably and has a high-application value.

Keywords Elevator energy consumption · Energy metering · Real-time monitoring · Digital isolation · ARM

76.1 Introduction

Elevators and escalators are the crucial elements that enable us to live, work, and go shopping, etc., conveniently and comfortably. For the sake of further urbanization in developing countries, an ageing population in many Western countries with the rising awareness of accessibility as well as a growing demand for convenience will lead to the installation of additional equipment. With the increasing prominence of climate warming and environmental issues, the implementation of energy-saving strategies has been attracting more and more attention with no exception

C. Lin (✉) · X. Wang · X. Wu · Z. Li
Guangzhou Academy of Special Equipment Inspection & Testing, 510663 Guangzhou, China
e-mail: linter0663@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_76

to elevators, which typically account for about 3–8 % of the overall electricity consumption of a building [1]. Researches show that the application of new saving technology such as permanent magnet synchronous driving, energy feedback, and intelligent control can help reduce the energy consumption (EC) of elevators [2]. The analysis of energy efficiency potentials indicates that considerable technical efficiency potentials exist for the elevators (more than 60 %) and escalators (around 30 %) in Europe [3]. Visibly, the energy-saving potential of elevators is substantial. The development of elevator energy consumption-monitoring device (ECMD) is the basis for EC monitoring, efficiency evaluation, and fault diagnosis [4, 5]; meanwhile, it is significant to adopt policy measures and corresponding retrofit scheme to promote the energy efficiency.

The ECMD involves the embedded processing, advanced sensor, and Internet of things (IOT). It samples the electrical signals of the elevator's main circuits and monitors the EC of the elevator real time. It is the basis to promote energy efficiency [6–8] by generating policy measures; however, few researches have been devoted to developing this device for monitoring the stand-by and running EC of the elevators. This chapter mainly focuses on the development of the ECMD. It includes the elevator's EC measuring principle, hardware design, software design, monitoring the stand-by and running EC of the elevator intelligently, and application.

The rest of this chapter is organized as follows: In Sect. 2, the basic principle of ECMD is introduced with the design of the hardware to be described in Sect. 3. In Sect. 4, we design the software and the EC-monitoring strategy. In Sect. 5, we explain the experimental setup, the calibration of the ECMD, and also the experiment results. Finally, we will summarize our results and arrive at some conclusions.

76.2 Basic Principle of ECMD

76.2.1 Structure of ECMD

In order to monitor the elevator's EC, the ECMD needs to measure the elevator's electric parameters, identify the running state, count the stand-by EC and running EC, and monitor whether the electric parameters of the elevator are abnormal. According to these requirements, the overall structure of ECMD is designed. The structure diagram is shown in Fig. 76.1 , which mainly includes the sensing unit, electric parameters acquisition unit, processing unit, communication unit, and power supply unit. The sensing unit is mainly composed of current sensors and voltage dividers. The sensing signal is conditioned first, then sampled and collected by the measurement chip. The results are isolated and sent to the processing unit. An ARM processor will calculate the measurement data and monitor the elevator's EC with the monitoring results sent to the monitoring center through the general packet radio service/wireless fidelity (GPRS/Wi-Fi) module.

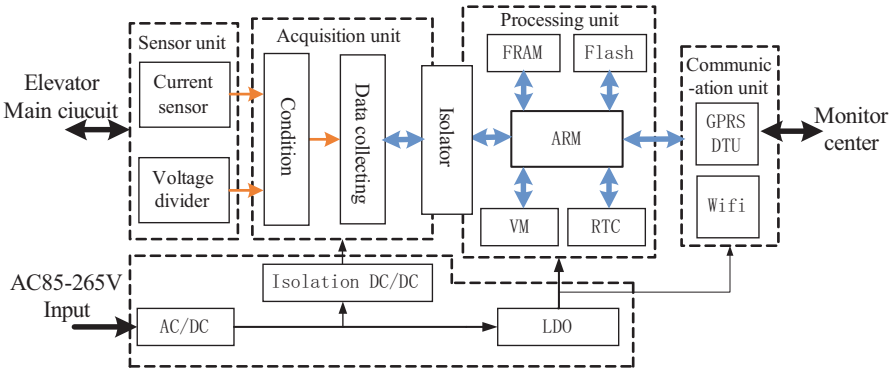


Fig. 76.1 Structure diagram of the elevator's ECMD ECMD elevator energy consumption-monitoring device

76.2.2 Basic Principle

The energy measurement is the basic function of the elevator's ECMD. The basic measurement principle of the electric parameters is discussed in this section. Generally speaking, the elevator adopts three-phase alternating current (AC) power supply. The EC monitoring of elevator is usually referred to the total active electric energy measurement and monitoring. Energy is the integral of instantaneous power $p(t)$, as follows:

$$e = \int p(t)\mathrm{d}t \bullet \tag{76.1}$$

Total active energy is the integral of the total active power, as follows:

$$E = \int P(t)\mathrm{d}t \bullet \tag{76.2}$$

The total active power is equal to the direct current (DC) component of the instantaneous power signal $p(t)$. In order to obtain the total active power, it is necessary to calculate the root mean square (RMS), which is characterized by the magnitude of an AC signal. As to an AC signal, the RMS value is the amount of DC required to produce an equivalent amount of power in the load. Mathematically, the RMS value of a continuous signal $f(t)$ can be defined as:

$$F_{\mathrm{rms}} = \sqrt{\frac{1}{t}\int_0^t f^2(t)\mathrm{d}t} \bullet \tag{76.3}$$

As to the discrete time sampling signals, RMS calculation involves squaring the signal, taking the average, and obtaining the square root, that is:

$$F_{\text{rms}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} f^2[n]}\bullet \tag{76.4}$$

Due to influences of the nonlinear load, a continuous signal $f(t)$ always contains harmonics [9], as follows:

$$f(t) = \sum_{k=1}^{\infty} F_k \sqrt{2}\sin(k\omega t + \phi_k). \tag{76.5}$$

So, Eq. (76.4) implies that the RMS calculation contains the contribution of all harmonics not only the fundamental. The first method is to low-pass filter the square of the input signal and take the square root of the result; that is:

$$\begin{aligned} f^2(t) = &\sum_{k=1}^{\infty} F_k^2 - \sum_{k=1}^{\infty} F_k^2 \cos(2k\omega t + \phi_k) \\ &+2\sum\nolimits_{\substack{k,m=1 \\ k\neq m}}^{\infty} 2\times F_k \times F_m \sin(k\omega t + \phi_k)\times\sin(m\omega t + \phi_m)\bullet \end{aligned} \tag{76.6}$$

Then, after the low-pass filter (LPF), the RMS value of $f(t)$ is obtained by

$$F_{\text{rms}} = \sqrt{\sum_{k=1}^{\infty} F_k^2}\bullet \tag{76.7}$$

The instantaneous power is defined as the rate of energy flow from source to load. As an AC system is supplied by the voltage $v(t)$ and the current $i(t)$, each of them contains harmonics, then:

$$v(t) = \sum_{k=1}^{\infty} V_k \sqrt{2}\sin(k\omega t + \varphi_k), \tag{76.8}$$

$$i(t) = \sum_{k=1}^{\infty} I_k \sqrt{2}\sin(k\omega t + \gamma_k), \tag{76.9}$$

where V_k, I_k are the RMS voltage and the current of each harmonic, respectively, φ_k and γ_k are the phase delays of each harmonic. The instantaneous power is obtained by:

$$\begin{aligned} p(t) &= v(t) \times i(t) \\ &= \sum_{k=1}^{\infty} V_k I_k \cos(\varphi_k - \gamma_k) - \sum_{k=1}^{\infty} V_k I_k \cos(2k\omega t + \varphi_k + \gamma_k) \\ &+ \sum_{\substack{k,m=1 \\ k \neq m}}^{\infty} V_k I_m \left\{ \cos\left[(k-m)\omega t + \varphi_k - \gamma_m\right] - \cos\left[(k+m)\omega t + \varphi_k + \gamma_m\right] \right\} \bullet \end{aligned} \tag{76.10}$$

So, the total active power is equal to the DC component of the instantaneous power signal $p(t)$ in Eq. (76.10), that is:

$$P(t) = \sum_{k=1}^{\infty} V_k I_k \cos(\varphi_k - \gamma_k) \bullet \tag{76.11}$$

As to the calculation of total active energy, the discrete time accumulation or summation is equivalent to the integration in continuous time following the description in Eq. (76.12).

$$E = \int P(t) \mathrm{d}t = \lim_{T \to 0} \left\{ \sum_{n=0}^{\infty} (P((n-1)T) + P(nT))T / 2 \right\} (n \geq 1), \tag{76.12}$$

where n is the discrete time sample number. T is the sample period.

76.3 Hardware Design

76.3.1 Power Supply Unit

The power supply unit includes AC/DC circuit and power supply circuit. Figure 76.2 is the scheme of the power supply unit. The AC/DC circuit consists of a WA10-220S12 power-supply module and a SPX1117M3-3.3 regulator chip. WA10-220S12 is used to convert the AC85~265-V AC input to DC 12-V output. SPX1117M3-3.3 regulates the DC input to supply the embedded system. B1203S-1W isolation power module converts the DC 12-V input to isolate DC 3.3-V output, which is the source of the electric acquisition unit. Its isolation voltage is 2.5 kV.

76.3.2 Electric Parameters Acquisition Unit

The electric parameters acquisition unit is the key component of the elevator's ECMD. It is based on multichannel Σ−Δ ADCs and specific DSP processing chip.

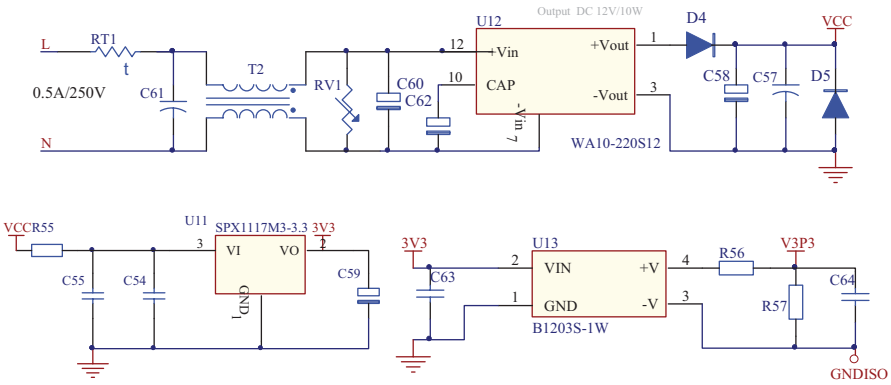


Fig. 76.2 Scheme of the power supply unit

Its peripheral circuits include current sensors, voltage dividers, filter networks, and digital isolation circuits, etc. The connection diagram of the electric parameters acquisition unit is shown in Fig. 76.3.

The current transformers are used to sense the main current of the elevator and transform to millivolt AC signals by burden resistors. The line voltages are attenuated by resistor divider networks. LPF networks are designed to attenuate the aliasing before they are supplied to the ADCs. The second-order Σ−Δ ADCs and specific DSP sample acquire the electric parameter and send it to the ARM processor upon isolation digitally.

Like other sampled systems, aliasing is an artifact of the electrical sampled systems regardless of the ADC architecture. It means frequency components above half of the sampling rate of the ADC (also known as the Nyquist frequency) are imaged or folded back down below half of the sampling frequency. In order to attenuate the high-frequency noise and prevent the distortion of the band of interest, an LPF is introduced. As to this sampled system, the sampling rate of the ADCs is 1024 kHz and the half is 512 kHz. To eliminate the effects of aliasing on the current sensors, 20 dB per decade attenuation of the filter is sufficient. A 1.43 kΩ resistance and a 22-nF capacity are selected to form a resistance-capacitance (RC) LPF with a corner frequency of 5 kHz for the attenuation to be sufficiently high.

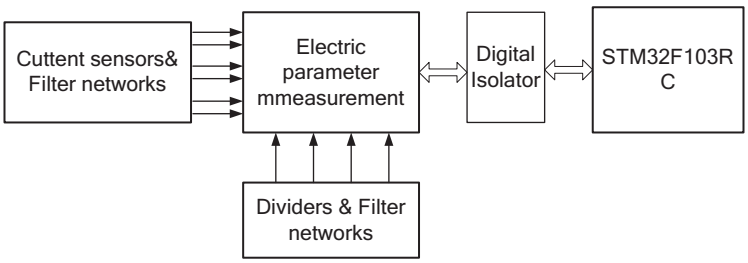


Fig. 76.3 Connection diagram of the electric parameters acquisition unit

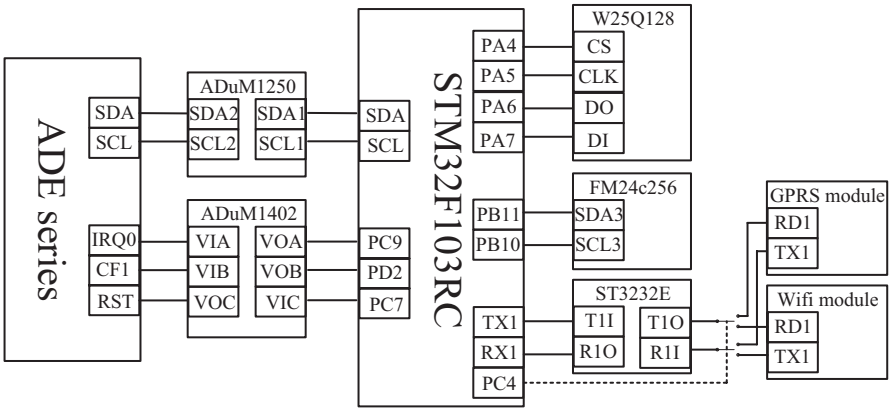


Fig. 76.4 External connection diagram of the ARM

76.3.3 Design of ARM Peripheral Circuit

STM32F103 ARM processor which is produced by STMicroelectronics (ST) is used as the main processor of the ECMD. It is based on the Cortex-M3 core with 32 Harvard micro structures, which can execute Thumb-2 instruction. Its frequency is 72 MHz, and it can achieve the processing speed of 1.2DMIPS/MHz. The external connection diagram of the ARM is shown in Fig. 76.4. Peripheral circuits include memory, data acquisition interface, and communication interface. The external memory includes W25Q128BV SPI Serial Flash and FM24C256 IIC ferroelectric memory. They are used to store the monitoring data and acquisition data, respectively. The data acquisition interface includes AduM1250 and AduM1402 magnetic-coupled isolators. The communication interface between the GPRS/Wi-Fi module and the ARM is the ST3232.

76.4 Software Design

76.4.1 Main Program

According to the function demand of ECMD, the main task of the software includes system initialization, data acquisition, data processing and storage, state recognition, intelligent monitoring, debugging console, and data communication, etc. The software design environment is uVison V4 using C language. The software flow diagram is shown in Fig. 76.5. The electric parameters of the elevator main circuit is collected and calculated to obtain the EC of elevator. The state-recognition judgment is mainly to determine whether the elevator is in stand-by or running state. The state of the elevator and the elevator's EC is monitored in a real-time time; and the stand-by EC and running EC are accumulated. Whenever they are abnormal, it

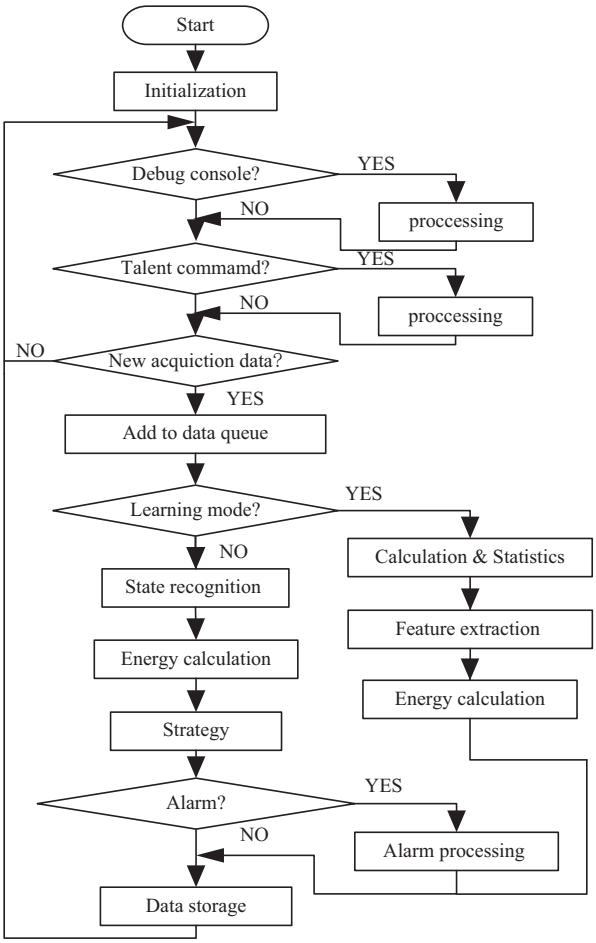


Fig. 76.5 Software flow diagram of the ECMD

will send alarm information to the monitoring center via the communication module intelligently.

76.4.2 EC Monitoring Strategy

The monitoring mode of the elevator's EC includes the self-learning mode and the real-time monitoring mode. In the learning mode, the monitoring strategy involves sampling the current of the elevator main circuit, extracting the stand-by feature signal, counting the stand-by EC and running EC, and estimating the anomaly threshold; and in the real-time monitoring mode, it distinguishes the running state from

stand-by mode according to the self-learned feature. It monitors the EC and judges whether the EC is more than the self-learned threshold intelligently. Whenever the elevator's EC or the current is abnormal, the alarm information will be sent to the monitoring center first. Second, the monitoring data can be used for remote fault diagnosis such as elevator low-insulation resistance, three-phase current unbalance, and mechanical failure of the traction system, etc.

76.5 Device Calibration and Experiment

The device debugging is mainly to verify the effectiveness of the acquisition hardware and the calibration procedure. The calibration command is sent from a PC to the ECMD by serial communication port. HIOKI3169 is used as a reference power meter. The debugging photo is shown in Fig. 76.6.

In order to further improve the accuracy of the device, it needs to calibrate the device referring to higher accuracy power meter. Referring to Fluke 6100 and calibrating by South China National Centre of Metrology, the results show that the measurement error of EC is +0.16%, the overall accuracy is better than 0.2%. The active power test results of B channel is shown in Table 76.1.

The parameters of the test elevator are shown in Table 76.2. After monitoring the elevator's EC for a month, the results are obtained as shown in Table 76.3. The movement EC is the running EC.

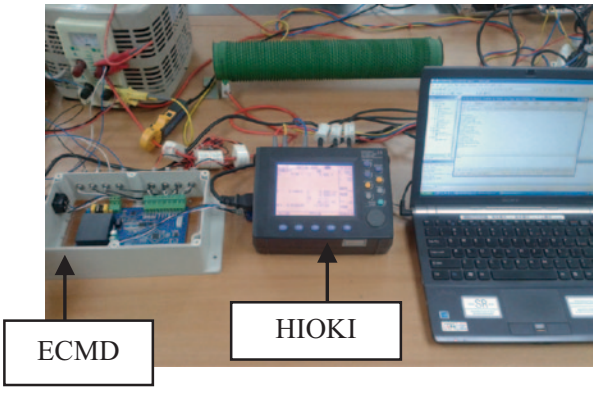


Fig. 76.6 Debugging photo of ECMD

Table 76.1 Active power calibration results of Channel B

Range	Indication value(W)	Reference value(W)	Error (%)
250 V/1 A	250.723	250.907	−0.07
250 V/2 A	500.718	500.831	−0.02
250 V/4 A	1001.08	1001.41	−0.03
250 V/20 A	5017.59	5020.43	−0.06

Table 76.2 Parameters of the test elevator

Item	Parameters
Brand	Guangri
Model	GVH1000–2S1.0
Rated load	1000 kg
Layer/ station	2/2
Control mode	VVVF
Transmission mode	Worm gear

VVVF variable voltage, variable frequency

Table 76.3 EC information statistics for a month

Item	Parameters
Movement EC (kWh)	40.95
Stand-by EC (kWh)	21.30
Stand-by AVG Power (kW)	0.1302

EC energy consumption, *AVG* average

76.6 Conclusion

This chapter proposes a scheme of online monitoring device of the elevator's EC, discusses the elevator's EC measuring principle, designs the hardware and software, and tests the device. Experiments show that, the EC measurement error is +0.16% and the overall accuracy is better than 0.2%. It may realize the monitoring of running EC and the stand-by EC intelligently; besides, it can provide technical support for remote EC monitoring and fault diagnosis. It meets the design requirements and has high application value.

Acknowledgments This R&D project is supported by the projects of Administration of Quality and Technology Supervision of Guangdong Province (2013PT05, 2010ZT02) and the science and technology projects of General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China (2012104016)

References

1. E4-Energy Efficient Elevators and Escalators. WP3 D3.2-Country Reports with the results of the monitoring campaign, 2009, Report elaborated for EC. http://www.e4project.eu/-articolo.asp?idarticolo=92. Accessed 21 April 2011.
2. Afonin V, Kruglikov O, Rodionov R. Energy consumption of elevator electric drives with different control systems. Russian Electr Eng. 2011;82(3):121–5.
3. Almeida AD, Hirzel S, Patrao C. Energy-efficient elevators and escalators in Europe: an analysis of energy efficiency potentials and policy measures. Energy Build. 2012;47(3):151–8.
4. Flores AQ, Carvalho JB, Cardoso AJM. Mechanical fault detection in an elevator by remote monitoring, Electrical Machines. Proceeding of 18th international conference on digital object identifier, IEEE, p. 1–5. 2008.
5. Hasan MZ, Fink R, Suyambu MR, Baskaran MK. Assessment and improvement of elevator controllers for energy efficiency. Proceeding of the 16th international conference on consumer electronics, IEEE, p. 1–8. 2012.
6. Lin CL, Wang XH, He B, Li ZX. Development of elevator condition monitoring device based on current feature. Autom Info Eng. 2013;34(4), p. 6–10.
7. Liu GX, Zhu HB, He RQ, Ke YH. Design of elevator real-time energy efficiency recorder and system. China Meas Test. 2012;38(6), p. 44–8 (In Chinese).
8. Lin CL, Wang XH, Wang WX. Study on current feature based FECMS for elevators in use. Proceeding of the 2013 international conference on mechatronic sciences, electric engineering and computer, IEEE, p. 750–3. 2013.
9. Wakileh GJ. Power systems harmonics. Beijing: China Machine Press; 2003. p. 6–14 (In Chinese).

Chapter 77
A Hierarchical Synthetic Aperture Radar Image Registration Method for Change Detection

Guangxue Wang, Yongchun Liu, Shirui Peng and Jiajun Zuo

Abstract The synthetic aperture radar (SAR) images obtained at different times often have both the global rigid deformation and the local elastic deformation. As a result, SAR image registration only based on the global rigid transformation model will induce an error, and the error will produce largely spurious results of change detection. In order to solve this problem, a hierarchical registration method is proposed in this chapter. It is a rough-to-fine registration procedure. In the rough registration stage, speeded-up robust reature (SURF) algorithm is first employed to extract the matching feature points from the testing image and the reference image. After that, a rigid registration method, based on the global affine transformation model, is implemented to correct the global deformation. In the fine registration stage, a template-match-based method is first adopted to extract a uniform spacing control point from the globally registered images. And then, a local elastic registration, based on thin-plate spline (TPS) transformation model, is applied to remove the local deformation. Experimental results show that the proposed approach can achieve higher registration accuracy, which lead to a better change-detection performance.

Keywords SAR · Change detection · Image registration

77.1 Introduction

Due to the all-time and all-weather capabilities of synthetic aperture radar (SAR) sensor, the change detection based on SAR images has wide potential applications in both civil and military fields [1, 2]; however, successful implementation of a change-detection analysis requires careful consideration of image registration methods because the misregistration will produce largely false alarms [3, 4]. The image

G. Wang (✉) · S. Peng · J. Zuo
Air Force Early Warning Academy, 430019 Wuhan, China
e-mail: wgxradar@163.com

Y. Liu
Beijing Institue of Technology, 100081 Beijing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_77

registration is the process of aligning two images of the same scene taken at different times, i.e., test image and reference image. Traditionally, this process is usually implemented based on a global rigid transformation model. Unfortunately, due to the impact of terrain relief and imaging track error, SAR images obtained at different times may have global rigid deformation and local elastic deformation simultaneously [5]. In this case, the misregistration will be inevitable as to the traditional method. To cope with this problem, we present a hierarchical registration method in this chapter. The method consists of two stages, i.e., the rough-image registration stage and the fine-image registration stage. In the stage of rough registration, a rigid registration approach based on SURF algorithm and affine transformation model is carried out to correct the global deformation; and in the stage of fine registration, a template-match-based method is first adopted to extract uniform spacing control points from globally registered images. After that, thin-plate spline (TPS), an effective transformation model of elastic image registration, is utilized to achieve local matching. The effectiveness of our method is validated by the experiment results.

The rest of this chapter is organized as follows: In Sect. 2, we state the rough stage of image registration. After that, the fine stage of image registration is described in Sect. 3. Then, in Sect. 4, the experimental results are presented. Finally, in Sect. 5, some conclusion remarks are drawn from this study.

77.2 Rough Stage of Image Registration

The goal of this stage is to get an initial registration result by removing the global deformation between the test SAR image and the reference SAR image. Its implementation process can be divided into two steps: The matching feature points extraction based on SURF algorithm and the global deformation correction based on affine transformation model.

77.2.1 SURF-Based Matching Feature Points Extraction

The matching feature points' extraction is the process to find feature point correspondence between the test and reference SAR images. It is a necessary part of the rough registration. In order to fulfill this task, we first have to detect feature points in the test and reference images, respectively, and construct a description for each of them. In the literature, a wide variety of feature point detectors and descriptors have been proposed. Among them, SURF algorithm is proved to be more reliable than other methods [6, 7]; thus it is also adopted in this chapter.

SURF algorithm is a novel scale- and rotation-invariant feature point detector and descriptor, presented by Bay in May 2006 [8]. The SURF detector can be constructed by using a scale space representation of an original image at different resolutions. It tries to find characteristic blob-like structures in an image independent

of its actual size. After detecting the feature points, the SURF descriptor describes each feature point by using a 64-dimensional vector. All these vectors include the local gradient direction and the magnitude information in a certain square neighborhood centered at the feature point. More details about SURF algorithm can be found in the original publications [8].

Once the feature point detection and description are fulfilled with SURF algorithm in both the test image and the reference image, the next step is to match each detected feature point in one image to the one in another image. In this chapter, a modified nearest-neighbor ratio criterion is used in the matching process.

Suppose ϕ_t and ϕ_r are two SURF feature point sets obtained, respectively, from the test and reference images, ψ_t and ψ_r are two SURF descriptor vector sets corresponding to ϕ_t and ϕ_r. Let $p_i^t \in \phi_t$, $p_i^r \in \phi_r$, $d_i^t \in \psi_t$, $d_i^r \in \psi_r$, d_i^t and d_i^r are two SURF descriptor vectors of p_i^t and p_i^r. If d_i^t and d_i^r satisfy:

$$\begin{cases} dist\left(d_i^t, d_j^r\right) = \min\limits_{d_k^r \in \phi_r} dist\left(d_i^t, d_k^r\right) = \min\limits_{d_l^t \in \phi_t} dist\left(d_l^t, d_j^r\right) \\ dist\left(d_i^t, d_j^r\right) \le \alpha \times \min\limits_{d_k^r \in \phi_r,\, k \ne j} dist\left(d_i^t, d_k^r\right) \\ dist\left(d_i^t, d_j^r\right) \le \alpha \times \min\limits_{d_l^t \in \phi_t,\, l \ne i} dist\left(d_l^t, d_j^r\right) \end{cases} \tag{77.1}$$

p_i^t and p_i^r can be regarded as a pair of matching feature points. Where $dist\left(d_i^t, d_j^r\right)$ represents the Euclidean distance between d_i^t and d_i^r, and $\alpha(\alpha \le 1)$ is the threshold. In this chapter, the value of α is set to 0.7.

77.2.2 Affine-Transformation Based Global Deformation Correction

In general, the global deformation between test and reference images can be modeled as an affine transformation. Given any point (x, y) in the test image $\mathbf{I}_t$ and its corresponding point (x', y') in the reference image $\mathbf{I}_r$, the affine transformation between $\mathbf{I}_t$ and $\mathbf{I}_r$ can be expressed as:

$$\begin{cases} x' = a_{11}x + a_{12}y + b_1 \\ y' = a_{21}x + a_{22}y + b_2 \end{cases}, \tag{77.2}$$

where $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ are the affine transformation parameter matrices.

Let $p_i = (x_i, y_i)$ and $q_i = (x_i', y_i'), i = 1, 2, \cdots N(N \ge 3)$ denote the corresponding feature points in the test and reference images, respectively. Substitute p_i and q_i into Eq. (77.2), thus the transformation parameter matrices $\mathbf{A}$ and $\mathbf{B}$ can be

estimated by the least-square method. After that, the global deformation between $\boldsymbol{I}_t$ and $\boldsymbol{I}_r$ is removed by applying Eq. (77.2) on $\boldsymbol{I}_t$ and a globally registered test image is obtained, which is denoted as $\boldsymbol{I}_g$.

77.3 Fine Stage of Image Registration

The goal of fine registration stage is to remove the local elastic deformation between the global registered test image $\boldsymbol{I}_g$ and the reference image $\boldsymbol{I}_r$. The implementation process of fine registration can be divided into two steps: The uniform spacing control points extraction based on template-match method and the local elastic registration based on TPS transformation model.

77.3.1 Uniform Spacing Control Points Extraction Based on Template-Match Method

Although a set of matching feature points have been extracted by SURF algorithm in rough registration stage, the number of points in it is often small and the space distribution of it is usually nonuniform; therefore, it cannot meet the demand of fine registration. In order to deal with this problem, in this stage, a template-match method is first employed to extract a uniform spacing control points from $\boldsymbol{I}_g$ and $\boldsymbol{I}_r$. Its construction process can be described in detail as follows:

First of all, many small subblocks according to a certain size are selected from $\boldsymbol{I}_g$. And the center of the subblocks is evenly distributed in $\boldsymbol{I}_g$. After that, as to each subblock, we adopt the normalized cross-correlation algorithm to search its most fitted subblock in $\boldsymbol{I}_r$ [9].

Given Ω_g and Ω_r are two subblocks belonging to $\boldsymbol{I}_g$ and $\boldsymbol{I}_r$, respectively, then the definition-normalized cross-correlation coefficient can be expressed as follows:

$$\gamma = \frac{\sum\left(f_g(i,j)-\overline{f}_g\right)\left(f_r(i,j)-\overline{f}_r\right)}{\sqrt{\sum\left(f_g(i,j)-\overline{f}_g\right)^2\sum\left(f_r(i,j)-\overline{f}_r\right)^2}}, \tag{77.3}$$

where $f_g(i,j)$, $f_r(i,j)$ is the gray value of any point (i,j) in Ω_g and Ω_r, $\overline{f}_g$ and $\overline{f}_r$ is the mean value of $f_g(i,j)$ and $f_r(i,j)$ within Ω_g and Ω_r. Based on Eq. (77.3), for each of the subblocks in $\boldsymbol{I}_g$, we can search its most fitted subblock in $\boldsymbol{I}_r$ by maximizing the value of γ. After that, for each couple of fitted subblocks, the centers of them are computed and marked as a pair of corresponding control points. Consequently, we can get $m \times n$ pairs of corresponding control points, where $m \times n$ is the number of subblocks in $\boldsymbol{I}_g$. Furthermore, because the control points in $\boldsymbol{I}_g(x,y)$ is evenly distributed, the distribution of control points in $\boldsymbol{I}_r$ is also close to be even.

77.3.2 Local Registration Based on TPS Transformation Model

In recent years, many different local elastic transformation models have been proposed, among which, TPS has been shown as an effective tool to implement local elastic registration, especially when the control points are evenly distributed across the whole image.

Let $u_i = (x_i, y_i)$ and $v_i = (x_i', y_i'), i = 1, 2, \cdots M$ denote the coordinates of corresponding control points in $\boldsymbol{I}_g$ and $\boldsymbol{I}_r$, respectively. The TPS transformation between $\boldsymbol{I}_g$ and $\boldsymbol{I}_r$ can be defined as:

$$\begin{cases} x' = a_0^x + a_1^x x + a_2^x y + \sum_{i=1}^{m} \omega_i^x \varepsilon\left(\left|u_i - (x, y)\right|\right) \\ y' = a_0^y + a_1^y x + a_2^y y + \sum_{i=1}^{m} \omega_i^y \varepsilon\left(\left|u_i - (x, y)\right|\right) \end{cases}, \tag{77.4}$$

where (x, y) is any point in $\boldsymbol{I}_g$, and (x', y') is its corresponding point in $\boldsymbol{I}_r$. $\varepsilon(|u_i - (x, y)|) = r^2 \log(r^2)$ is the bi-harmonic equation, in which r is the distance between u_i and (x, y). a_0^x, a_1^x, a_2^x, ω_i^x, a_0^y, a_1^y, a_2^y, and ω_i^y are unknown parameters of TPS transformation model. Besides Eq. (77.4), there are three more constraints on ω_i^x and ω_i^y:

$$\sum_{i=1}^{m} \omega_i^x = 0, \sum_{i=1}^{m} \omega_i^y = 0, \tag{77.5}$$

$$\sum_{i=1}^{m} x\omega_i^x = 0, \sum_{i=1}^{m} x\omega_i^y = 0, \tag{77.6}$$

$$\sum_{i=1}^{m} y\omega_i^x = 0, \sum_{i=1}^{m} y\omega_i^y = 0\cdot \tag{77.7}$$

Based on Eq. (77.4) ~ Eq. (77.7), a_0^x, a_1^x, a_2^x, ω_i^x, a_0^y, a_1^y, a_2^y, and ω_i^y can be calculated by replacing (x, y) and (x', y') with $u_i = (x_i, y_i)$ and $v_i = (x_i', y_i')$. After that, the local elastic deformation between $\boldsymbol{I}_g$ and $\boldsymbol{I}_r$ is removed by applying Eq. (77.4) on $\boldsymbol{I}_g$, and a finally registered image is obtained.

77.4 Experimental Results

To verify effectiveness of our proposed algorithm, a group of experiments is implemented in this section. Figures 77.1a and 77.1b are the test and reference SAR images used in the experiments, which are taken at an area of little forest. As it

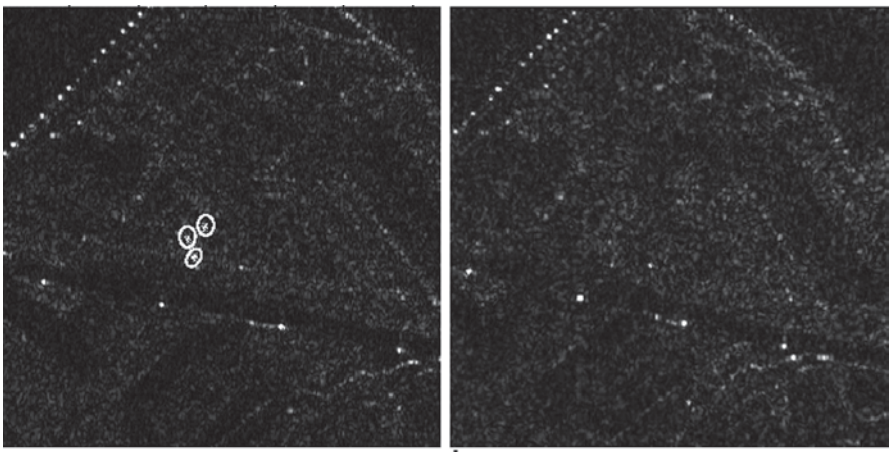


Fig. 77.1 Two SAR images. **a** Test image. **b** Reference image

is shown, there are three corner reflectors placed in Fig. 77.1a and removed in Fig. 77.1b, which are labeled with white circles in Fig. 77.1a.

Figure 77.2 shows the results of image registration, whereas Fig. 77.2a is the result by traditional method, which is only based on a global rigid transformation model. Figure 77.2b is the image registration result by our proposed method.

As we all know, as to the change detection, the higher accuracy of registration can reduce the false alarms, and the lower accuracy of registration will produce more false alarms; therefore, in this chapter, in order to compare the registration accuracy of Fig. 77.2a and Fig. 77.2b, experiments on change detection are carried

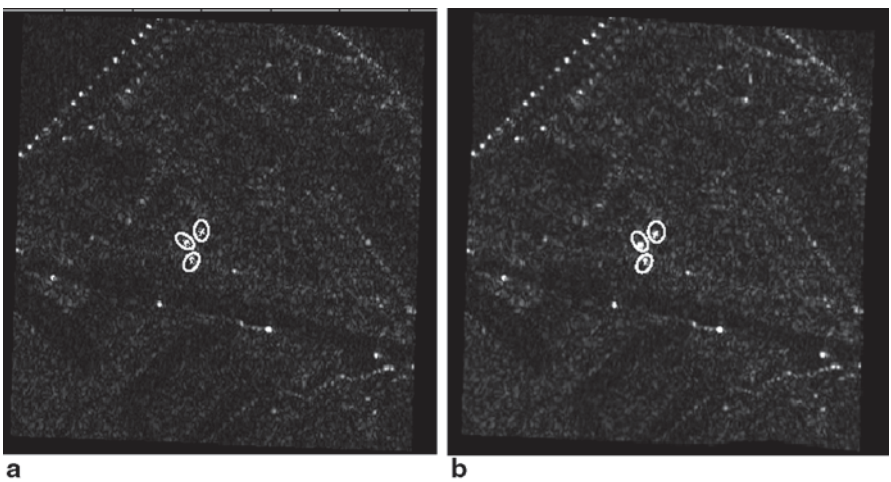


Fig. 77.2 Result of image registration. **a** Registration result obtained in rough stage. **b** Registration result obtained in fine stage

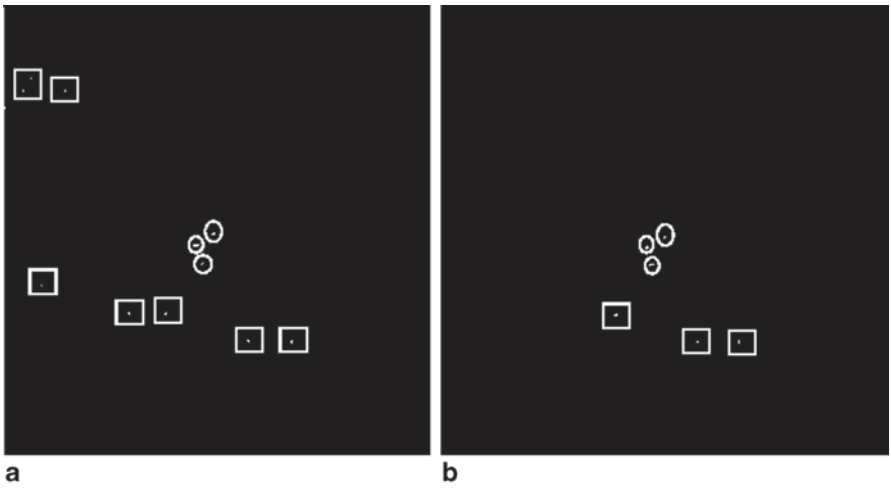


Fig. 77.3 Change detection results. **a** the result of change detection by comparing Fig. 77.2a and Fig. 77.1b. **b** the result of change detection based on Fig. 77.2b and Fig. 77.1b

out on this basis. Details about the algorithm of change detection can be found in the relative publications [10]. The change detection results are shown in Fig. 77.3, where the white points in the square are detected as false alarms. Obviously, the number of false alarms in Fig. 77.3b is much less than that in Fig. 77.3a; so we can conclude that our proposed method has a better registration performance than the traditional method.

77.5 Conclusion

A hierarchical SAR image registration method for change detection method is proposed in this chapter. The method consists of two stages: the rough-image registration stage and the fine-image registration stage. In the rough-image registration stage, a rigid registration approach based on SURF algorithm and affine transformation model is carried out to correct the global deformation. In the fine-image registration stage, the uniform spacing control points are first extracted from the globally registered images by the template-match method. Then, a local elastic registration, based on TPS transformation model, is applied to remove the local deformation. The TPS-based image registration (TPS) is utilized to achieve local matching. The effectiveness of our method is validated by the experiment results.

Acknowledgments In this chapter, the research was sponsored by the Nature Science Foundation of China (Project No. 61302194).

References

1. Ulander LMH, Flood B, Frolind P. Change detection of vehicle-sized targets in forest concealment using VHF- and UHF-band SAR [J]. IEEE Aerosp Electron Syst Mag. 2011;26(7):30–6.
2. Giustarini L, Hostache R, Matgen P. A change detection approach to flood mapping in urban area using TerraSAR-X [J]. IEEE Trans GRSS. 2013;51(4):2417–30.
3. Dirk F, Peter HN. Misregistraion errors in change detection algorithms and how to avoid them [C]. IEEE International conference on image processing, IEEE. 2005;pp. 438–441.
4. Theiler J, Wohlberg B. Local coregistration adjustment for anomalous change detection [J]. IEEE Trans GRSS. 2012;50(8):3107–16.
5. Xie H, Pierce LP. Mutual information based registration of SAR images [C]. 2003 International journal of computer vision, IEEE. 2003;pp. 4028–4031.
6. Salgian A. Object recognition using local descriptors: a comparison [C]. Berlin: ISVC, Spriger; 2006.pp. 709–17.
7. Mikolajczyk K, Tuytelaars T, Schmid C. A comparison of affine region detectors. Int J Comput Vision. 2005;65(1):43–72.
8. Bay H, Ess A. Speeded-up robust features (SURF). Comput Vision Image Underst. 2008;110(3):346–59.
9. Mattoccia A, Tombari F, Stefano L. Fast full-Search equivalent template matching by enhanced bounded correlation. IEEE Trans Image Process. 2008;17(4):528–38.
10. Ranney KI, Soumekh M. Sigbnal subspace change detection in average multilook SAR imagery [J]. 2006;IEEE Trans GRSS. 44(1):201–13.

Chapter 78
Fast Beamforming Based on SPSA for Mobile Satellite Antenna

Lanlan Zhang and Yingguang Hao

Abstract As to the phased array antenna in the mobile satellite reception, a novel fast beamforming algorithm is presented based on simultaneous perturbation stochastic approximation (SPSA). Compared to previous algorithms, the proposed approach estimates the gradient of the received power simply relative to the phase difference between two adjacent phase shifters because the difference remains the same in a uniform array. All phase shifters are updated simultaneously according to the gradient. As the digital phase shifters are assigned from a discrete set, there is an error between the theoretical and practical values. The error is regarded as a random perturbation when the gradient is estimated. A large number of simulation experiments verify that the proposed algorithm is effective and efficient in both the elevation beam scan and the two-dimensional beam scan and that it provides a good tracking performance to obtain the optimal signal.

Keywords Beamforming · Digital phase shifter · Phased array antenna · SPSA

78.1 Introduction

As to the mobile phased array antennas, it has become more important to use fast beamforming algorithms [1]. An electronic beamforming algorithm is always utilized to reach the accurate direction of the satellite after the mechanical coarse alignment procedure [2]. The simultaneous perturbation stochastic approximation (SPSA) algorithm estimates the gradient of the received power to update the phase shifters via an iterative process [1, 3, 4].

The gradient estimation plays a core role in SPSA. There is a method which estimates the gradient simultaneously by perturbing all phase shifters at the same time;

L. Zhang (✉) · Y. Hao
Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, 116033 Dalian, China
e-mail: zhangll50@mail.dlut.edu.cn

Y. Hao
e-mail: yghao@dlut.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_78

however, it is likely to diverge in some cases [1]. Taking into account the locations of array elements, an improved method gets the gradient in the positive direction rather than the full isotropy. It converges quickly, but may not achieve the optimum because of the superposition error with the complicated *N*-dimensional computation [2].

In this chapter, a fast beamforming algorithm based on the algorithm in the second references is presented. Digital phase shifters are chosen instead of analog ones which are easily impacted by the mutative environment variables [5, 6]. Thus, the phase difference for adjacent antenna elements can be considered reasonably the same regardless of the phase shifters' different attributes in different situations. The new method significantly reduces the complexity of computation and converges quickly. In addition, it is applied to two-dimensional beam scan; consequently, the performance of the tracking system is further improved.

The rest of this chapter is organized as follows. In Sect. 78.2, the antenna structure and pattern in this chapter are introduced clearly. In Sect. 78.3, the essential analysis for this algorithm is shown particularly. In Sect. 78.4, the experimental results with different methods are presented. The conclusions are drawn in Sect. 78.5.

78.2 Antenna Structure and Pattern

Figure 78.1 shows a planar antenna array placed in H-plane. It has a uniform distribution with M*N antenna elements. As the antenna array is far enough away from the target satellite, it is reasonable to assume the radiated pattern of each element is fully isotropic and that the excitation signals are uniform [7]. The normalized function of the antenna pattern is

$$\left|F(\phi,\theta)\right| = \left| \frac{\sin(\frac{M\times be}{2})}{M\times\sin(\frac{be}{2})} \frac{\sin(\frac{N\times ba}{2})}{N\times\sin(\frac{ba}{2})} \right| \tag{78.1}$$

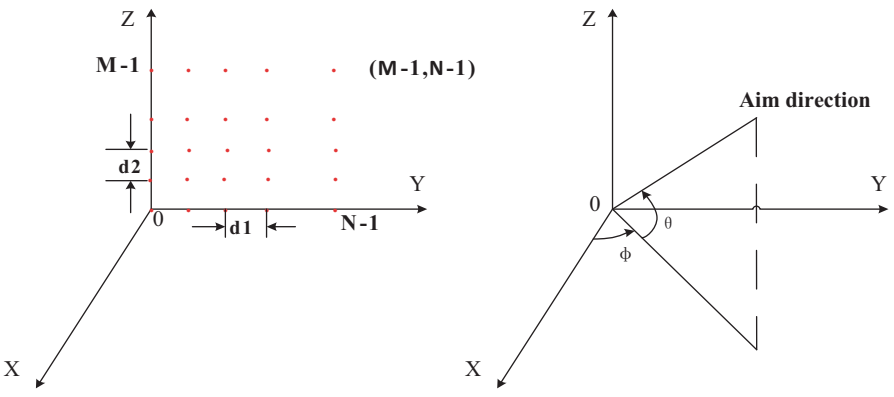


Fig. 78.1 Diagram of a planar phase array antenna

where $be = \frac{2\pi d_2}{\lambda}(\sin\theta - \sin\theta_0)$, $ba = \frac{2\pi d_1}{\lambda}(\sin\phi\cos\theta - \sin\phi_0\cos\theta_0)$. The parameters θ_0 and ϕ_0 are respectively the elevation and the azimuth of the antenna normal. The parameter λ denotes the wavelength.

78.3 Fast Beamforming Algorithm

Supposing the antenna normal deviated from the aim direction is set to $(\Delta\phi, \Delta\theta)$, the fine tuning is achieved by regulating the phase shifters to reduce it. The purpose of this algorithm is to receive the maximum power of the satellite signal if it is possible. The measured power without any prior knowledge is described as the equation $P(\phi,\theta) = |F(\phi,\theta)|^2$. As is known to all, the optimum methods for maximizing a nonlinear function can be settled though an iterative process which uses gradient for fast approximation [3].

78.3.1 Elevation Scan in H-Plane

The planar antenna array is equivalent to a column linear array in this case. The phase difference Δe is the only factor to be obtained. The processing steps of the new algorithm are as follows:

1. Let the variable $\Delta e_1 = 0$, which denotes the phase difference at the first iteration. All the phase shifters are initialized to be zero.
2. The subscript k represents the iteration number. At the k^{th} iteration, the value of Δe is

$$\Delta e_{k+1} = \Delta e_k + a_k g_k(\Delta e_k) \tag{78.2}$$

where $a_k = \frac{a}{(A+k)^\beta}$ is the step size. When the step size increases, the convergence speed raises, but the steady state behaves differently. It is necessary to find a proper value for the parameter a to perform effectively [1, 3, 4]. The parameter $g_k(\Delta e_k)$ denotes the gradient of the received power relative to Δe at the k^{th} iteration, which is estimated by

$$g_k(\Delta e_k) = \frac{P(\Delta e_k + b_k D\Delta\theta\xi_k + c_k\Delta_k) - P(\Delta e_k - b_k D\Delta\theta\xi_k - c_k\Delta_k)}{2(b_k D\Delta\theta\xi_k + c_k\Delta_k)} \tag{78.3}$$

where $b_k = \frac{b}{(k+1)^\gamma}$, $c_k = \frac{c}{(k+1)^\gamma}$. $D = d_2$ denotes the distance between two adjacent antenna elements. The parameter $\Delta\theta$ is the angle jitter error. Both ξ_k and Δ_k are the random perturbation scalars from a Bernoulli distribution with equal probability of each value which is +1 or −1. In practice, the small value $c_k\Delta_k$ is considered as a

stochastic perturbation factor to reduce the affection caused by the quantization of the digital phase shifters.

3. The iterative process stops until the received power reaches a threshold in practice; otherwise, it returns to step 2 to continue. The threshold is determined by the user requirements. For example, when it comes to the mobile satellite TV reception, the threshold can be the proper value which satisfies the data rate to obtain good performance for watching.

78.3.2 Two-Dimensional Beam Scan in Space

Compared to the mechanical scan, an electronic scan has two obvious advantages: fast beamforming and inertialess scanning. Therefore, it tracks the target satellite more quickly and accurately in question. The process of the simultaneous elevation and azimuth scan is shown as below:

As mentioned above, there are also three steps to follow; however, a little change is made here. The scalar Δe is replaced by the vector $\Delta s = [\Delta a, \Delta e]^T$, where Δa is the phase difference between two adjacent antenna elements in a row. The phase difference between the unit in the location (i, k) and the origin is $i * \Delta e + k * \Delta a$. The approach computes Δs for the iterative process. In Eq. 78.3, g_k is the gradient of the received power relative to Δa and Δe, estimated respectively as follows: Keeping Δe constant, use Eq. 78.3 to compute the gradient relative to Δa; do the same with Δa.

78.4 Experimental Results and Analysis

In this section, the previous algorithm in the second references and the proposed algorithm in this chapter are applied to the elevation scan to compare the rate of convergence; besides, the elevation and the azimuth scan are achieved simultaneously by the new algorithm as well.

In the simulation, an 8*8 planar array antenna receiving signal at the frequency of 12 GHz around Ku-band is chosen. The distance d1 and d2 are set to be λ/2. Suppose the antenna normal deviated from the satellite position is 4.5°. Regarding the initial antenna normal as the coordinate datum, the elevation ϕ_0 and the azimuth θ_0 of the target position are both 4.5°.

78.4.1 Elevation Scan in H-Plane

To conduct the experiment, the distribution of the satellite signal in H-plane is simulated firstly. Assuming the angle is positive above the coordinate datum and negative on the contrary, the signal model is

$$P(\Delta e) = \left| \frac{\sin[M(e_0 - \Delta e)/2]}{M \sin[(e_0 - \Delta e)/2]} \right|^2 \tag{78.4}$$

where $e0 = \frac{2\pi d_2}{\lambda} \sin(rm * \theta_0)$ is the spatial phase difference. rm denotes the random variable from a Bernoulli distribution with uniform probability of each value which is 1 or −1.

Set the internal algorithm parameters to be the proper values: a = 0.025, b = 0.02, c = 0.01, A = 0.1, $\beta = 0.062$, $\gamma = 0.101$, $\Delta\theta = 0.00002$. The values are determined by a mass of simulation experiments which have been performed in the second references. Assuming the normalized power reaching 0.95 seems to meet the convergence condition, the convergence diagrams of different algorithms are respectively shown in Fig. 78.2 and Fig. 78.3. As to each scenario, 100 runs are performed independently.

The previous method takes 25 iterations on average to converge, while the new one only takes nine iterations. It indicates that the new algorithm tracks the target much faster. Supposing there is a mistake when the number of the iterations surpasses 500, the error rate of the new method is 0.

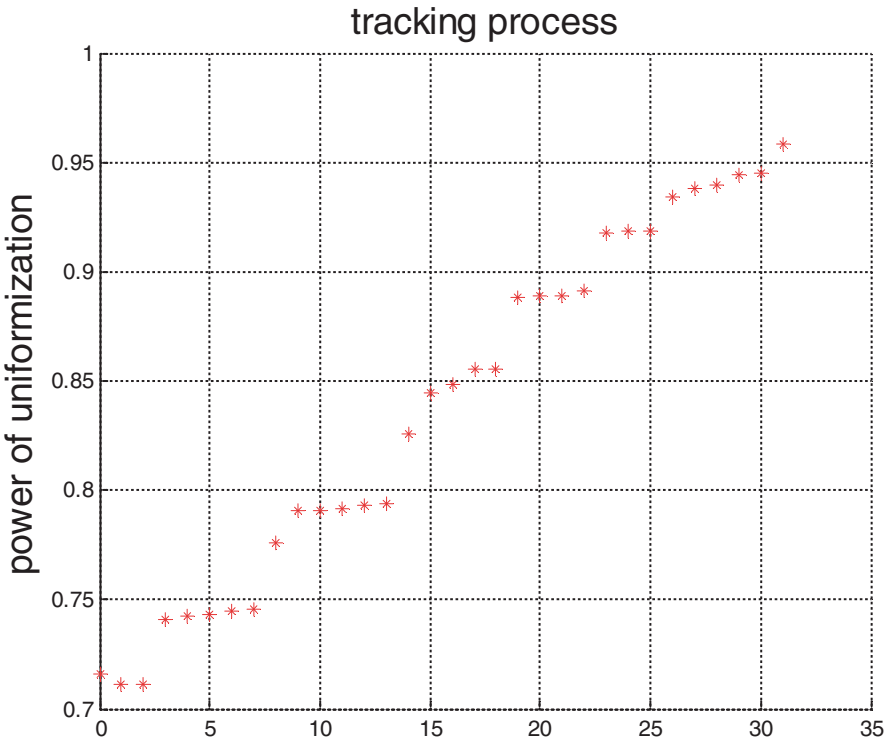


Fig. 78.2 Convergence of the previous algorithm

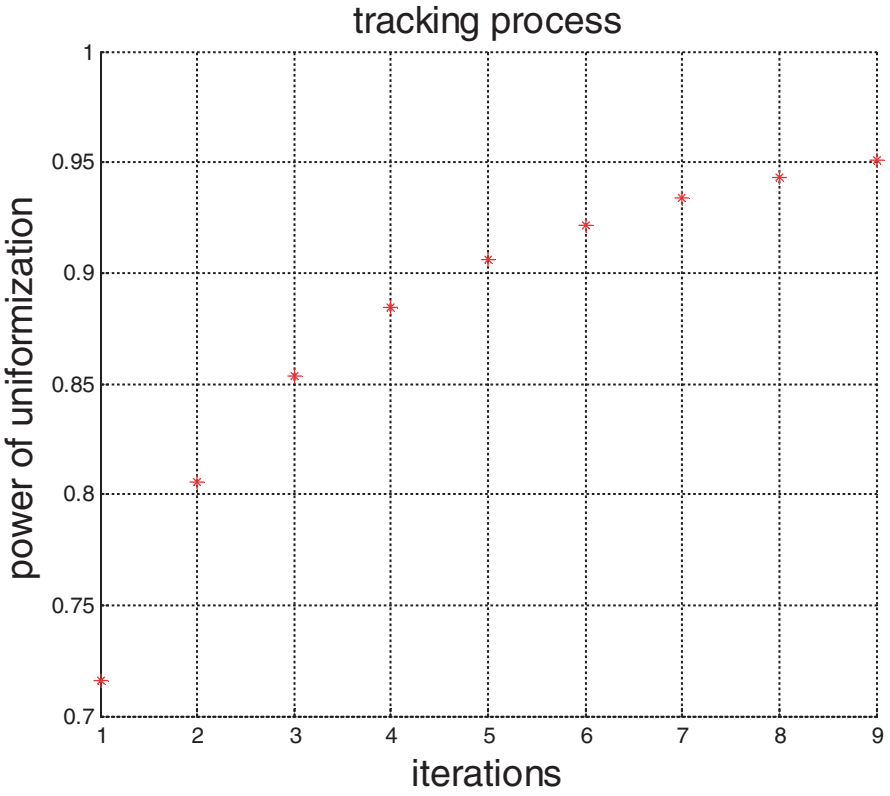


Fig. 78.3 Convergence of the new algorithm

78.4.2 Two-Dimensional Beam Scan

Similar to the procedure of the elevation scan, a signal model is built firstly about the phase difference $(\Delta a, \Delta e)$. The function is

$$P(\Delta a, \Delta e) = \left| \frac{\sin[M(e_0 - \Delta e)/2]}{M \sin[(e_0 - \Delta e)/2]} \bullet \frac{\sin[N(a_0 - \Delta a)/2]}{N \sin[(a_0 - \Delta a)/2]} \right|^2 \tag{78.5}$$

where $a_0 = \frac{2\pi d_1}{\lambda} \sin(rm1 * \phi_0) \cos(rm2 * \theta_0)$ and $e_0 = \frac{2\pi d_2}{\lambda} \sin(rm * \theta_0)$ are respectively the spatial phase difference in E-plane and H-plane. The random variables rm1 and rm2 have the same attributes as rm. The diagram of the signal model in space is shown in Fig. 78.4.

Figure 78.5 shows the convergence of the two-dimensional beam scan, which uses the proposed method with the same internal parameters.

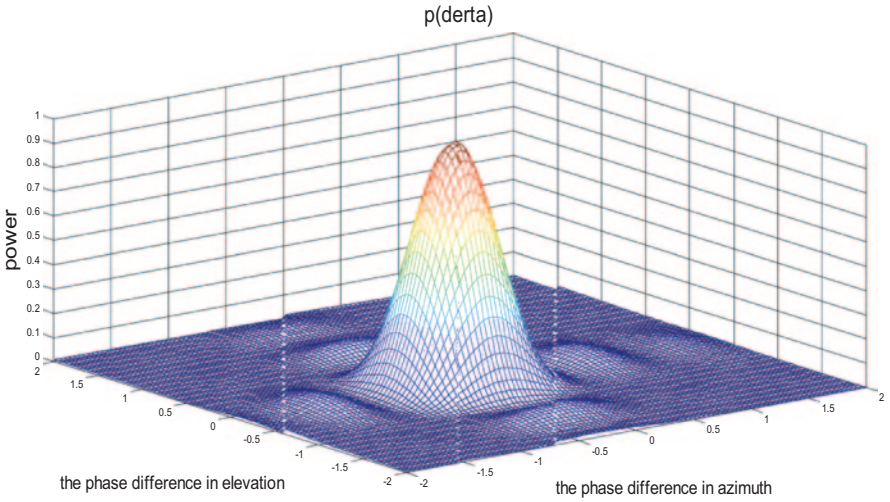


Fig. 78.4 Diagram of the signal model in space

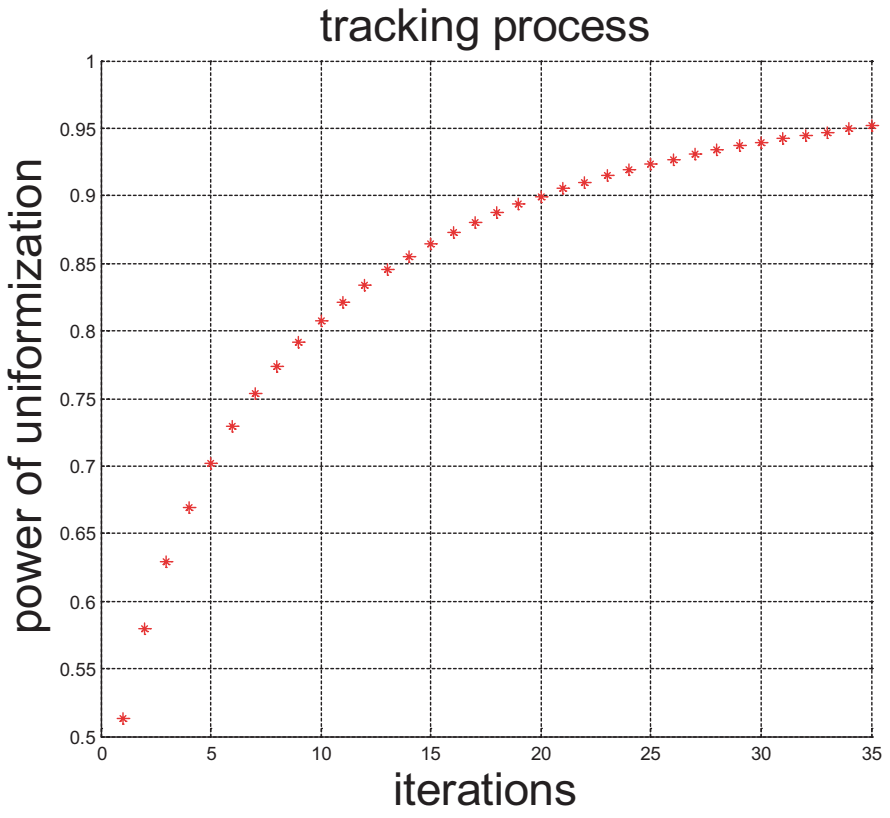


Fig. 78.5 Convergence of the new method in two-dimensional beam scan

With 100 runs performed independently, there are 35 iterations on average to converge. Compared to the elevation beam scan, it needs more iterations because of more impact factors; besides, there are two more times updating the phase shifters during each gradient estimation. The convergence speed and accuracy get much better by reasonably assigning the parameter a. For example, the experimental results show it needs only 11 times to converge when a = 0.05. The two-dimensional beam scan takes full advantage of the electronic scan, making the tracking much more effective and efficient.

78.5 Conclusion

This chapter presents a fast beamforming algorithm with constraints to obtain the optimal signal. Using digital phase shifters with a stable performance, a reasonable perturbation is made to estimate the gradient relative to the phase difference, which reduces the computation dimensions from N to 1. The proposed algorithm improves the convergence speed and accuracy. Furthermore, it is innovatively applied to two-dimensional beam scan. In contrast with a hybrid tracking scheme which achieves the azimuth scan mechanically, it obviously performs much better; however, as the algorithm neglects the positional errors of the antenna elements and the drift errors of the digital phase shifters, the received power may not be able to achieve the optimum in practice. The algorithm will be further improved to reduce actual errors in the future work.

References

1. Fakharzadeh M, Jamali SH, Safavi-Naeini S, Mousavi P, Narimani K. Fast stochastic beamforming for mobile phased array antennas. Antennas and Propagation Society International Symposium, 2007 IEEE, IEEE, p. 1945–48. 2007.
2. Zhang FG, Jia WM, Jin W, Yao ML. Beamforming algorithm based on SPSA for mobile satellite receiver. Electron Lett. 2012;48(22):1379–80.
3. Fakharzadeh M, Jamali SH, Narimani K, Mousavi P. Zero-knowledge beamforming for mobile satellite phased array antenna. Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th, IEEE, p. 1–5. 2008.
4. Fakharzadeh M, Jamali SH, Mousavi P, Safavi-Naeini S. Fast beamforming for mobile satellite receiver phased arrays: theory and experiment . IEEE Trans Antennas Propag. 2008;57 (6):1645–54.
5. Azadegan R. A Ku-band planar antenna array for mobile satellite TV reception with linear polarization. IEEE Trans Antennas Propag. 2010;58(6):2097–101.
6. Gou YX, Fu JH, Fang XQ, Wu Q. Two-dimensions beam scanning reflectarray antenna. Antennas and Propagation Society International Symposium(APSURSI), 2013 IEEE, IEEE, p. 1674–75. 2013.
7. Dybdal RB, SooHoo KM. Narrow beamwidth satellite antenna pointing and tracking. Antennas and propagation (APSURSI), 2011 IEEE International Symposium on, IEEE, p. 2012–15. 2011.

Chapter 79
A Reference Architecture of Data Acquisition and Signal Processing of a Distributed Fiber-Optic Sensor System

Qiwu Wang, Gangyang Nan, Jihua Sun, Jiangfeng Song, Rui Guo and Jian Ma

Abstract Distributed fiber-optic sensors require fast data acquisition. Large amount of sensors on the network require enormous area and bandwidth of memory on the devices. Besides, the data transmission scheme of distributed fiber-optic sensors introduces more complexity in signal processing. The chapter contributes an embedded solution to the problems confronted by the digital system of distributed fiber-optic sensor systems. A reference architecture is introduced in this chapter, which provides a template to resolve the problems with a field-programmable gate array (FPGA) chip. A datapath with a finite-state machine is implemented to control the analog-to-digital (A/D) converter and achieve specified algorithms; a soft-core processor is implemented to achieve functions of computation and human–machine interface. This architecture shrinks the size of memory and circuit board. The chapter also provides a detailed design to demonstrate customized instantiation of the architecture.

Keywords Distributed fiber-optic sensor · Data acquisition · Signal processing

79.1 Instruction

Distributed fiber-optic sensors are capable of monitoring large amounts of physical indices with a single fiber. These sensors have advantages of lightweight, immunity to electromagnetic interference, strong corrosion resistance, and capability of continuous real-time remote sensing [1]. Recent research shows that Raman distributed temperature sensors are able to measure the temperature with the accuracy of ±1 °C with the range of 80–300 °C [2].

The large throughput of distributed fiber-optic sensor systems is challenging circuit designers. Thousands of sensors lead to enormous memory demands for data

Q. Wang (✉) · G. Nan · J. Sun · J. Song · R. Guo · J. Ma
Laser Research Institute, Shandong Academy of Science, 28789 Jinan, China
e-mail: qiwu@outlook.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_79

storage. The data transmission scheme of distributed fiber-optic sensors makes the online signal processing more complex. The chapter contributes a low-cost field-programmable gate array (FPGA) solution that data acquisition, signal processing, and program of human–machine interface (HMI) are implemented on a single chip. The cost is largely reduced.

Section 2 describes challenges confronted by the digital design of distributed fiber-optic sensors; Sect. 3 presents a reference architecture for the application. Section 4 proposes a case study which demonstrates a customized instantiation of the architecture. Section 5 contains a discussion of the above design.

79.2 Challenges of the Digital Design

The data acquisition and signal processing of distributed fiber-optic sensors inherits general characteristics of other digital systems. However, characteristics of the fiber system bring more explicit challenges:

1. Data transmission scheme of distributed fiber-optic sensors complicates the signal processing. When an optic pulse is triggered, signal beams are scattered or reflected from sensing points which are allocated with space intervals between each other. Therefore, the data from each point arrive at analog-to-digital converter (ADC) following a time sequence. The sample sets can be noted as

$$\{S_1[n], S_2[n] \ldots S_m[n] \ldots\} \tag{79.1}$$

where m represents a series of sensing points, and n represents a sample array triggered by n^{th} optic pulse. If the previous sample fetched from ADC is $S_m[n]$, then the sample on the next rising edge will be $S_{m+1}[n]$, rather than $S_m[n+1]$. Therefore, the traditional schemes which hold the previous sample with registers do not work in this case; and most of electronic design automation (EDA) build-in blocks, like digital filters and Fourier transforms, cannot be used in this application.

2. High space resolution and signal processing require high throughput and large memory. The space resolution of distributed fiber-optic sensors is proportional to its sampling frequency [3], which implies the throughput of signal processing should be high enough to support the corresponding sampling frequency. Besides, the processing data of thousands of sensors require huge memory space.

A common laboratory solution is that a data acquisition device fetches all the data, then transmit the data to a terminal PC and process with software [4–6]. This method is easy to apply and helps users focusing on other research issues, but it is too cumbersome for field applications. DITEST STAR-R series designed by Omnisens provide a compact and intelligent solution. The device has advanced user interfaces and automatic configuration for different sensors. DITEST STAR-R series have superior computation capability and a 160-GB hard disk for data storage [7]. They are used in

applications like strain and cracking monitoring [8]. But these devices are very expensive. A low-cost system can enlarge the application of distributed fiber-optic sensors.

79.3 Reference Architecture

To improve compactness and economy of distributed fiber-optic sensor systems, a reference architecture based on low-price FPGA is introduced in this section. The designer can use this architecture as a template and modify it to achieve customized requirements.

The reference architecture of the overall system consists of a pack of digital circuit, which controls dataflow and implements algorithms, as well as a soft-core processor to implement HMI, transmit data, and carry out part of the algorithm. Figure 79.1 shows the overall reference architecture. The digital circuit is composed of blocks of memory, a datapath, counters, and a finite state machine (FSM). The soft-core processor is, at least, composed of a CPU, memory controller, as well as flash memory, and related interface. Other peripherals are application dependent.

79.3.1 Datapath and Finite State Machine

The datapath consists of registers, random-access memory (RAM) blocks, and arithmetic logic units (ALU). In this architecture, RAM blocks are used to cache the previous value for specific algorithms. The outputs of RAM are fed back to the inputs of ALUs. Thus, the previous data $S_m[n-i]$ and the current data $S_m[n]$ can be used as operands for the ALUs.

The FSM controls the operation of ADC, RAM address, and the operation of ALU. In this case, writing and reading address are controlled independently. Note that the on-chip RAM, the quickest memory available, still has one clock cycle delay to response the change of reading address. Thus, an important feature of FSM is fetching the data in advance to guarantee the correct operand is fed into the ALU when a new sample arrives.

The corresponding instruction set is also flexible to implement different processing algorithms. With the purpose of simplifying the digital design, a generic

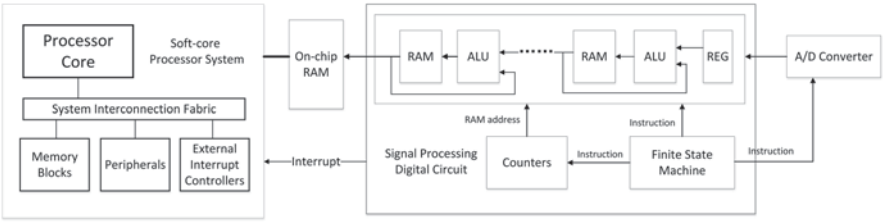


Fig. 79.1 Overall system architecture

Fig. 79.2 Instruction set

Table 79.1 Instruction set example

Blocks	Interpretation of instruction
ADC ctrl	"1-": ADC enable; "-1": ADC Standby
ALU ctrl	"000": NOT; "001": AND; "010": XOR; "011": OR; "100": Add; "110": Subtract; "111": Bypass
RAM enable ctrl	"1-": RAM read enable; "-1": RAM write enable
Counter ctrl	"1–": counter enable; " –1": clear the count; "110": count increment; "100": count decrement

ADC analog to digital converter, *ALU* arithmetic logic units, *RAM* random-access memory (The symbol "-" means "don't care.")

instruction is given in Fig. 79.2, and an example of the interpretation of the instructions is listed in Table 79.1. Based on this template, the designer can justify their instructions according to explicit requirements.

79.3.2 Soft-Core Processor

A soft-core processor is a processor core described in hardware description languages (HDL) and could be targeted to FPGA or application-specific integrated circuit (ASIC) chips. The processors provided by vendors are configurable, meaning that a user can add or remove features on a system-by-system basis to meet the performance or price goals [9]. Compared to traditional hard processor, the soft-core processor has advantages in flexibility and configurability.

In this architecture, the processor interfaces the datapath with a RAM block. The write operation is controlled by the FSM and read operation is controlled by the processor. When processed data are stored in the RAM, an interrupt signal is sent. Then the processor will fetch the data from the interface memory, process them, display, or update them.

79.4 A Case Study: Distributed Temperature Fiber-Optic Sensors

A customized design is introduced below to demonstrate an instantiation of the reference architecture. An embedded device is required to fetch data from distributed temperature fiber-optic sensors and calculate average value over at least 60,000

times. The resolution of ADC is 10 bits and the sampling frequency is 50 MHz. The data of 5000 sampling points need to be acquired and averaged.

In this measurement scheme, a measuring period is triggered by an optic pulse. During a measuring period, scattered light beams which carry the temperature information are transmitted to a photodetector. Then the light beams are converted to electrical signals and received by ADC. It implies that the average value is calculated based on the sample sets which are triggered by more than 60,000 optic pulses.

The interval between each optic pulse is 100 µs. For the efficiency, the average values are calculated over 65,536 times, by which the division operation could be replaced by shift operation. Thus, the average value of *m* sensing points is represented as

$$y_m = \frac{1}{65536} \sum_{n=0}^{65535} S_m[n] \tag{79.2}$$

79.4.1 Design of Datapath and FSM

The digital system is implemented on an FPGA chip of Altera Cyclone IV E. The schematic of datapath with FSM is shown in Fig. 79.3. The datapath is composed of an adder block, shift logic, RAM blocks, and registers. The adder block has the adding function as well as directly passing the input values.

The register after the ADC synchronizes samples as well as expand data width from 10 to 26 bits before feeding into the adder block. This function is to avoid data overflow during the sum calculation. When is feeding to the adder, the FSM will fetch the sum $\sum_{0}^{n-1} S_m[n]$ in advance, so as to make sure both of them will be fed to the adder at the same clock event. Then, the new sum $\sum_{0}^{n-1} S_m[n]$ is stored in the RAM1; when the sum reaches 65,536 times, the accumulated data will be transmitted to shift logic. Then, the shift logic right-shift the data 16 bits to achieve the division of 65,536. The data width expanded to 16 bits at the output of RAM2 to fit the data width of the processor.

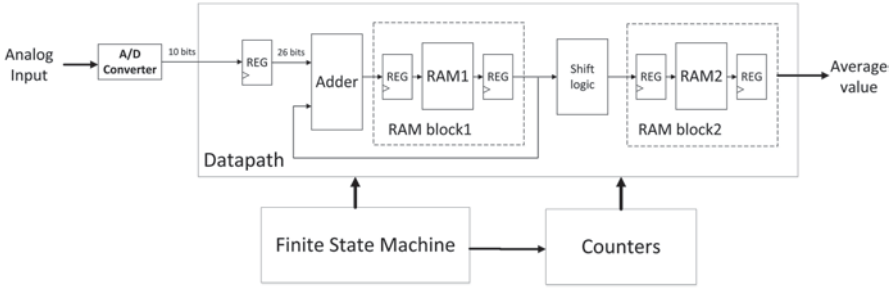


Fig. 79.3 The datapath with finite state machine and counters

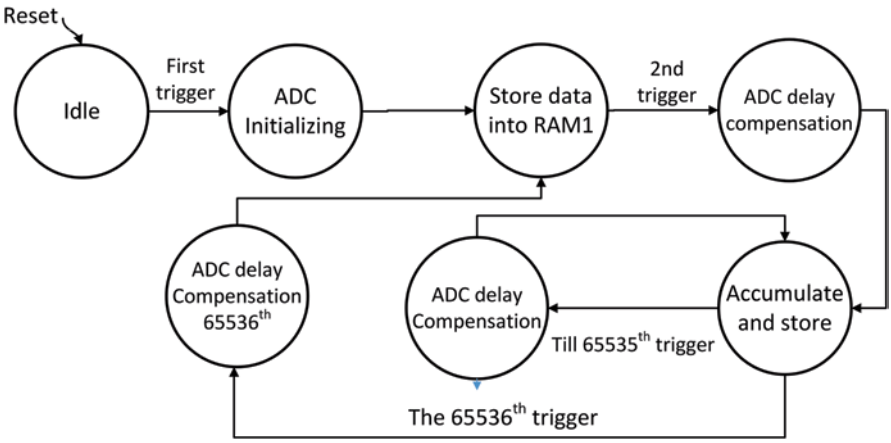


Fig. 79.4 State transition graph of the FSM

The FSM controls the operation of ADC and data flow in the datapath. A mealy machine is adopted in this case. The Fig. 79.4 shows the state-transition graph. The state "Idle" is the default state in which no operation is made. When the first trigger comes, the state will transfer to "ADC Initializing," in which the ADC is enabled. The sample set triggered by first pulse are outputted after the ADC time delay and stored in the RAM1. Then, the second trigger comes. But due to the delay of ADC, the first few samples are still triggered by the first pulse. These data should be stored where the previous sample set is stored. Thus, in state "ADC delay compensation" same clock cycles of delay is compensated for the address operation. Then, the sample set $S_m[2]$ are arrived, added with $S_m[1]$, and then stored in RAM1. This process iterates until the 65536^{th} trigger, the state back to preparing for the new first trigger. At the same time, the sum values are right-shifted before fed into RAM 2 and an interrupt is send to the processor.

79.4.2 Design of Instruction Set

The design of instruction set refers to the template given in Sect. 3. One modification is that an average counter is implemented to count the times of triggers. In this case, the instruction set has a width of 20. Table 79.2 shows a description of the instruction set. Note that the read address of RAM2 is controlled by the soft processor.

79.4.3 The Nios II Soft-Core Processor

The Nios II core provided by Altera Corporation is adopted. In this case, Nios II/f, which is the fastest core in Nios II family, is implemented. The data cache is set as

Table 79.2 The instruction set and related description

Bits	Name	Description
0–1	ADC control	Enable and standby the ADC
2	Input enable	Enable the input of the datapath
3	Adder control	Switch between "sum" or "bypass"
4–5	RAM1 enable	Enable read/write port of RAM1
6–7	RAM2 enable	Enable read/write port of RAM2
8–10	RAM1 read address ctrl	Counts read address of RAM1
11–13	RAM1 write address ctrl	Counts write address of RAM1
14–16	RAM2 write address ctrl	Counts write address of RAM2
17–19	Average counter ctrl	Counts the time of trigger signal

ADC analog to digital converter, *RAM* random-access memory

8 Kb and the data cache line size is 4 byte. On-chip RAM of processor has a width of 16 and depth of 204,800 byte. The software start fetches the data from RAM2 when an interrupt is received; then average data is interpreted into temperature information and displayed on the liquid-crystal display (LCD) screen.

79.4.4 Simulation Result and Resource Occupation

The function of the digital circuit is simulated in SignalTab and the results are plotted in MATLAB. The simulation used ideal sinusoidal waves of 20 KHz after each trigger. The Fig. 79.5 shows the input and address signals. As stated above, the read addresses are fetched in advance comparing to the write addresses. Figure 79.6

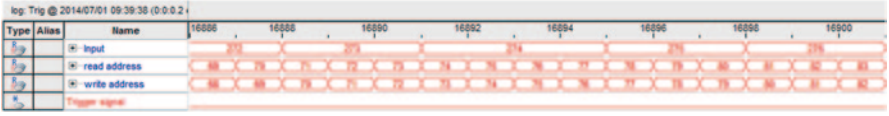


Fig. 79.5 Input and addresses in SignalTab simulation

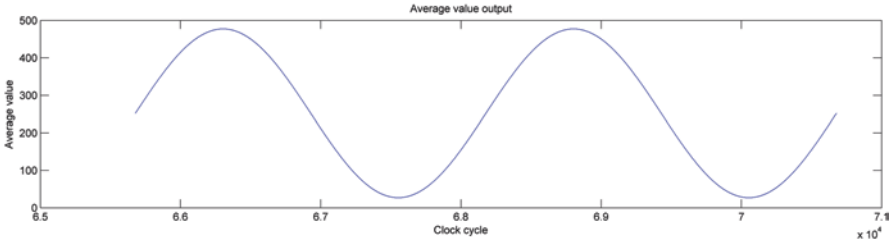


Fig. 79.6 SignalTab simulation plotted in MATLAB

shows the output average value. The wave of average values retains the shape of input signal. After compiling, overall the system only occupies less than 4000 logic elements and 2,000,000 memory bits. The maximum clock frequency of digital circuit is up to 141 MHz. The low resource consumption implies that the system can be implemented on a low-cost FPGA chip.

79.5 Conclusion

A reference architecture of data acquisition and signal processing of distributed optic-fiber sensor is introduced in this chapter. The tasks of data acquisition, signal processing, and HMI are implemented on a single FPGA chip. The reference architecture provides a template of the digital design and a case study demonstrates a customized instantiation. Therefore, the development cost could be largely reduced.

Acknowledgments The research in this paper is sponsored by the College and University Innovation Project of Jinan (Project No.201401209).

References

1. Bao T, Wang J, Yao Y. A fiber optic sensor for detecting and monitoring cracks in concrete structures [J]. Sci China Tech Sci. 2010;53(11):3045–50 (In Chinese).
2. Fukuzawa T, Shida H, Oishi K. Performance improvements in Raman distributed temperature sensor [J]. Photonic Sensors. 2013;3(4):314–9.
3. Cao L. Research on distributed fiber optic temperature measurement and data processing [D]. Hefei: Heifei University of Technology. 2006.
4. Zhang C. FBG intelligent perimeter security method and key technology research [D]. Wuhan: Wuhan University of Technology. 2013.
5. Cong J. Study on temperature monitoring by distributed optical fible and the intelligent feedback for hydraulic quasi-mass concrete structure [D]. Wuhan: Wuhan University. 2013.
6. Xie S. Polarization properties of Brillouin scattering and interferometry in optical fibers and their applications on distributed fiber sensing [D]. Beijing: Tsinghua University. 2013.
7. Omnisens Corporation. http://www.omnisens.com/ditest/32-Aboutus.php. Accessed 26 June 2014.
8. Mao J. Research on application of distributed optical fiber sensor in structural strain and cracking monitoring [D]. Zhejiang: Zhejiang University. 2012.
9. Altera Corporation. Nios II processor reference handbook. 2010.

Chapter 80
Soft-Sensing of Coal Ash Content Based on Dual-Energy Γ-Rays

Dong Cheng, Fengwen Shen, Zhaosheng Teng, Siyu Guo and He Wen

Abstract To improve the measurement accuracy of coal ash content, this chapter proposes an intelligent soft-sensing method based on dual-energy γ-rays by using the functional link neural network. This method takes ^{241}Am and ^{137}Cs as the resources of low-energy and medium-energy γ-rays, respectively, and uses the γ count of the detector as auxiliary variable. The coal ash content is measured and verified after accomplishing the model of soft-sensing with the functional link neural network that is optimized by chaos algorithm. The result shows that the method of functional link neural network forecasting based on chaos optimization algorithm has higher accuracy and excellent extensive capability than conventional nuclear technology. The intelligent soft-sensing based on functional link neural network with optimized chaos algorithm forecasting model is of measurement accuracy, and the most error and mean error between the soft-sensing values and real values are 0.9 and 0.7%.

Keywords Coal ash content · Soft-sensing · Dual-energy γ-rays · Chaos optimization algorithm · Functional link neural network

80.1 Introduction

Ash is the main economic indicator of the coal industry analysis, whose content is defined as the amount of residue after burning under prescribed conditions [1]. The ash value is obtained by means of sampling technologies associated with the conventional chemical analysis. Although traditional methods have been used to determine coal properties, they are laborious and time-consuming procedures which are not suitable for a real-time methodology [2].

To achieve a better and faster measurement value of coal ash content, nuclear technologies have been adopted widely now. Nuclear technologies (collectively known as inelastic-scattering technology of low-energy γ-ray, low-energy γ-ray penetration, dual-energy γ-ray penetration, etc.) have been divided into the backscattering and penetration technologies [3–5]. These conventional nuclear technologies

D. Cheng (✉) · F. Shen · Z. Teng · S. Guo · H. Wen
College of Electrical and Information Engineering, Hunan University, 410082 Changsha, China
e-mail: kaixinguiace@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_80

are based on linear empirical formula of detected γ-ray intensity and ash to measure the coal ash content by parameter calibration. In order to get a better coal ash measurement with little sample, this chapter proposes a novel intelligent soft-sensing based on functional link neural network (FLNN) with chaos optimization algorithm forecasting model.

Because of the nonlinear signal processing and learning capability, artificial neural networks have become a powerful tool for many complex applications including functional approximation, nonlinear system identification and control, pattern recognition and classification, and optimization [6–8]. The FLNN is basically a flat net and the need of the hidden layer is removed, and the backpropagation (BP) learning algorithm [9] used in this network becomes very simple. The functional expansion effectively increases the dimensionality of the input vector and the hyperplanes generated by the FLNN provides greater discrimination capability in the input pattern space [10–12]. Chaos variable's traversal property ensures that it can get approximate solution with extremely good precision even if the optimization calculation time is limited [13, 14]. We believe the method proposed in this chapter will promote the development of coal industry because of the higher precision and advantages than the conventional methods. In Sect. 2, the basic theory of dual-energy γ-rays is introduced in brief. The choice of auxiliary variables in coal ash soft sensor and the model for FLNN and the chaos optimization algorithm in FLNN are present in Sects. 3 and 4. Simulation results are provided in Sect. 5. The conclusions are drawn in Sect. 6.

80.2 Basic Theory of Dual-Energy Γ-Rays

80.2.1 Basic Theory of Measurement

Coal is combined with combustible organic and noncombustible minerals. The direct measurement of the coal content is very difficult to realize, so the ash can be measured to reflect the noncombustible mineral content and as an indicator of the quality assessment of coal.

Coal is regarded as a mixture of two kinds of atomic number elements generally. One is a representative with C, relatively low atomic number elements, referred to as low-Z elements. The others are representative with Si and Al, referred to as high-Z elements, and account for about 51% of ash quality. Therefore, as long as the determination of high-Z elements' quality of coal sample, and multiplied by two, is the value of coal ash approximately, $A = 2C_z$.

The measurement framework of coal ash system by using dual-energy γ-ray is shown in Fig. 80.1.

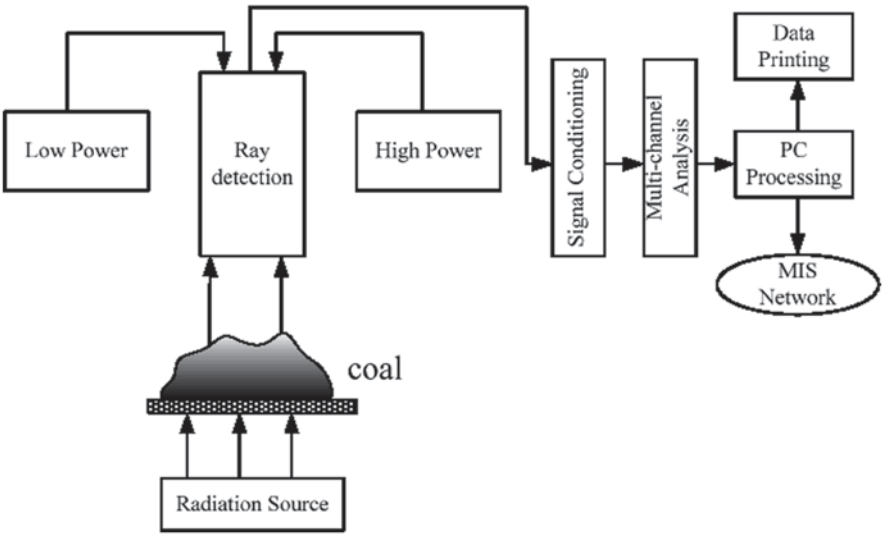


Fig. 80.1 Measurement structure of coal ash content based on dual-energy γ-rays

80.2.2 The Measurement Model of Ash

In dual-energy γ-rays, one is a low-energy γ-ray radiation and the other is a middle-energy γ-ray radiation. Both of γ-ray beam collimated passes through the coal, it will attenuates exponentially by γ-ray of narrow beam transmission. And the attenuation law of low-energy γ-ray is

$$I_1 = I_0 e^{-u_L \rho d}. \tag{80.1}$$

Here, I_0 is the number of low-energy γ when interval detector test of nothing to a certain time and I is the count after coal absorption. μ_L is the coal mass attenuation coefficient of the γ-ray. ρd is the mass thickness of transmitted coal. ρ is bulk density and d is thickness of coal. Therefore,

$$\ln(I_1/I_0) = -u_L \rho d \tag{80.2}$$

and so that

$$\ln(I_1/I_0) = -\left[\mu_{C(L)} + C_z\left(\mu_{Z(L)} - \mu_{C(L)}\right)\right]\rho d \tag{80.3}$$

$$\ln(J_1/J_0) = -\left[\mu_{C(H)} + A\left(\mu_{Z(H)} - \mu_{Z(H)}\right)\right]\rho d \tag{80.4}$$

In the formula, J_0 is the number of low-energy γ when interval detector test of nothing to a certain time and J is the count after coal absorption. $\mu_{Z(L)}$ and $\mu_{Z(H)}$ are the mass attenuation coefficients of the reaction of high-Z elements with low-energy γ ray and with middle-energy γ ray, cm^2/g. $\mu_{C(L)}$ and $\mu_{C(H)}$ are the mass attenuation coefficients of the reaction of low-Z elements with low-energy γ ray and middle-energy γ ray, cm^2/g. Because of the same coal seam, the ρd is the same. As to middle-energy ray, $\mu_{C(H)}$ and $\mu_{Z(H)}$ can be considered as the same. $\mu_{C(L)}$, $\mu_{Z(L)}$, $\mu_{Z(H)}$ can be determined by experiment, C_Z can be obtained easily by the following formula:

$$C_z = \frac{\ln(I_1/I_0)}{\ln(J_1/J_0)} \cdot \frac{\mu_{Z(H)}}{\mu_{Z(H)} - \mu_{Z(H)}} + \frac{\mu_{C(L)}}{\mu_{Z(L)} - \mu_{Z(L)}} \tag{80.5}$$

Because of this, ash can be considered approximately as two times as the quality elements of high-Z element; the ash of coal is measured by dual-energy γ-rays, $A \approx 2C_Z$.

80.3 Model for Coal Ash Soft Sensor

80.3.1 The Choice of Auxiliary Variables in Coal Ash Soft Sensor

Choose the ^{241}Am as low-energy γ source, and ^{137}Cs as the middle-energy γ source to measure the ash content by dual-energy γ-rays generally. Through the NaI detector, the γ intensity can be counted by signal processing and transmission spectroscopy analysis. Then the ash content can be calculated and predicted by the intelligent ash content model of soft-sensing based on chaos optimization algorithm in FLNN structure.

80.3.2 Model for FLNN

FLNN has some features of functional approximation, nonlinear system identification and control, pattern recognition and classification, and optimization. Through the nonlinear transformation of function extension, converting each input component X_k corresponding linearly independent function $f_1(x_k), f_2(x_k), \ldots f_n(x_k)$. Therefore, new information presentation space is extended to enable a single network which can be distinguished from complex objects.

For the prediction of m-dimensional phase space to nonlinear time series consisting of $X=(x_1, x_2, \ldots x_m)$, suppose the number of conventional forecasting methods are $K(K>2)$. These are $f_1(x_i), \ldots, f_j(x_i), \ldots, f_K(x_i) (i=1, 2, \ldots, m, j=1, 2, \ldots, K)$, and as the inputs $u_1=f_1(x_i), \ldots, u_i=f_i(x_i), \ldots, u_K=f_K(x_i)$. After a series of extensions available

to obtain the function input values and form the actual inputs of FLNN. Considering the mean square value, it can perform function expansion through cross trigonometric functions as brief as the other orthogonal basis approached functions.

For a predefined function expansion order of *S,* the actual input *v* of the neurons is $\{u_k, \{\cos(s\pi u_k), sin(s\pi u_k)\}\}(k=1, 2, \ldots, K; s=1, 2, \ldots, S)$. It will get extra auxiliary inputs as many as $N=2\times S\times K$. The activation function of single neuron is Sigmoid, and the output *Y* of FLANN can be established as follows:

$$Y = 1/\left(1+e^{-z}\right) \tag{80.6}$$

Here, $z=\theta_0+y,\theta_0$ is the threshold of neuron. And

$$y = \sum_{m=1}^{K+N} w_m v_m \tag{80.7}$$

In the formula, v_m is obtained by the input u_n original function expansion. The FLNN prediction model formed is shown in Fig. 80.2.

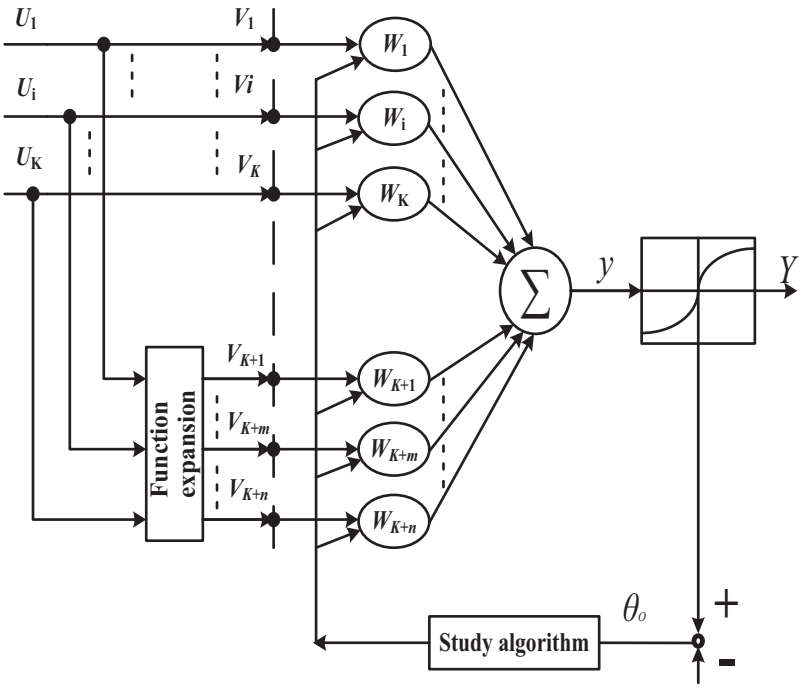


Fig. 80.2 Forecasting model of functional link neural network

80.4 Chaos Optimization Algorithm in FLNN

80.4.1 Necessary and Sufficient Conditions for Fitting

Suppose there are K original inputs $u_i(i=1, \ldots, K)$ and become $N+K$ components: $v1_i$, $v2_i$, …, $v_{(N+K)i}$ after function expansion and $N+K+1$ weight components: w_0, w_1,…, $w_{(N+K)}$. In order to $Y_i=x_i$, so $z_i=\ln[x_i/(1-xi)]$. x_i and z_i correspond with each other. Because of $\Sigma w_i \cdot v_i-\theta=z_i$, if $w_0=-\theta$, we get the formula $w_0+\Sigma w_i \cdot v_{ii}=z_i$.

As to complex process of nonlinear time series $X=(x_1, x_2, \ldots, x_m)$, so

$$\begin{bmatrix} 1 & v_{11} & \cdots & v(N+K)1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & v_{1m} & \cdots & v(N+K)m \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N+K} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} \tag{80.8}$$

If the weight $W=[w_0, w_1, \ldots, w(_N+_K)]$ is obtained, the question of fitting nonlinear time series can be solved easily.

80.4.2 Chaos Optimization Algorithm with the Weight Parameters

As to FLNN, the ultimate goal is to optimize the minimum error function, the general error function is

$$Error = \frac{1}{2}\sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2, \tag{80.9}$$

where N is the total number of samples, $\hat{Y}_i$ is the prediction output of neural network, and Yi is the real output of samples. FLNN is a practical solution of nonlinear problems, such as:

$$\begin{cases} \min Error(W,b) \\ W,b \in [-1,1] \end{cases}. \tag{80.10}$$

Here, W is the weight matrix of neural networks, b is threshold matrix.

Chaos optimization algorithm must select the appropriate chaotic model firstly, and propose the Eq. 80.11 as the map model.

$$x_{n+1} = \sin(2/x_n), x_n \in (-1,1) \tag{80.11}$$

Chaos optimization algorithm is to solve the nonlinear programming problem such as Eq. 80.10, and it has combined with crude searching and fine searching in the actual calculation process.

80.5 The Simulation and Experiment

In order to predict the intelligent ash soft measurement model based on the chaos optimization algorithm in FLNN, the experiment uses 20 sample data of one mine as the basis for the training data, to train through ash intelligent soft measurement model. Coal samples are taken from multiple coal sources with diversity and representation.

An NaI scintillation detector which detects the γ intensity or counting by dual-energy γ-rays of ^{241}Am and ^{137}Cs and uses chemical methods to measure the ash true value of coal particles.

It has been shown that the soft sensor ash predictive value is consistent with the true value and the error is smaller after training through in Fig. 80.3 and in Fig. 80.4. The most error and mean error between the soft-sensing values and real values is 0.9% and 0.7%. And the root mean square error (RMSE) of soft-sensing values is 0.59%.

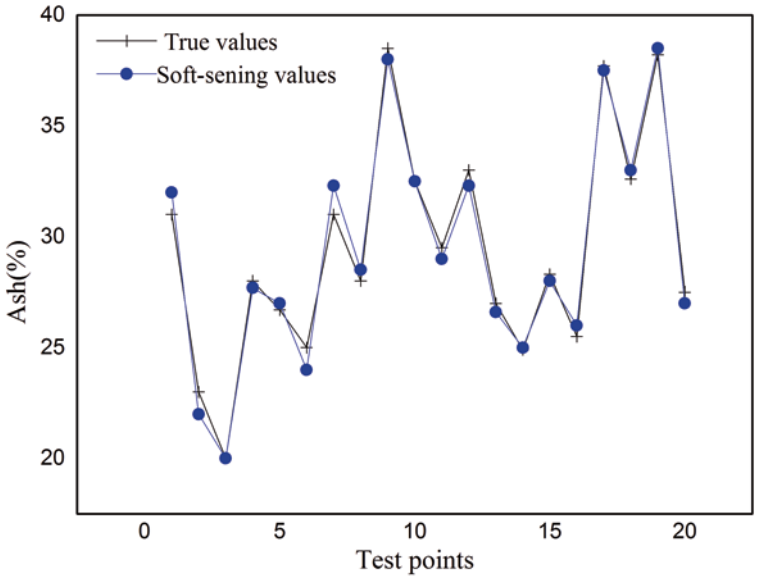


Fig. 80.3 Comparison of soft-sensing values with real ash content values

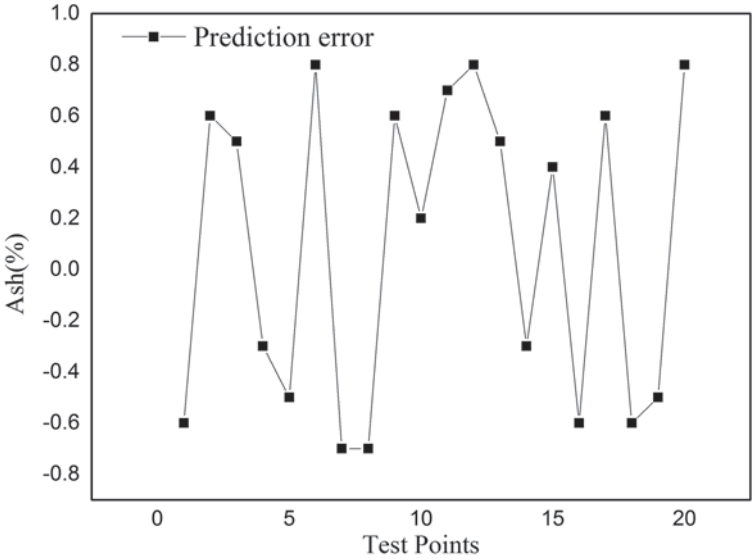


Fig. 80.4 Prediction error of soft-sensing values

80.6 Conclusion

In the analysis of neural network function chain fitting on the basis of necessary and sufficient conditions, it is proposed that the FLNN prediction model based on chaos optimization algorithm is to be combined with the weight of FLNN.

The FLNN forecasting method based on chaos optimization algorithm has been used in coal ash measurement and established a coal ash intelligent soft-sensing based on dual-energy γ-rays. The samples are predicted by soft-sensing and test through training and validation sample data.

A brief overview of higher forecasting accuracy and extensive capability is presented in this study, as well as the most significant applications using both simulation and experimental verifications. All of that prove the method has good measurement accuracy and advantages compared with the conventional measurements.

Acknowledgments This work was supported in part by the National Natural Science Foundation of China under Grant 61370014 and Grant 51377049.

References

1. Lim CS, Abernethy DA. On-line coal analysis using fast neutron-induced gamma-rays. Appl radiat isot. 2005;63(5):697–704.
2. Kim DW. Application of near infrared diffuse reflectance spectroscopy for on-line measurement of coal properties. Korean J Chem Eng. 2009;26(2):489–95.

3. Takeuchi S. Low-lying states in Mg 32 studied by proton inelastic scattering. Phys Rev C. 2009;79(5):314–9.
4. Suster D, et al. Accuracy of dual energy X-ray absorptiometry (DXA), weight and P2 back fat to predict whole body and carcass composition in pigs within and across experiments. Livest prod sci. 2003;84(3):231–42.
5. Takahashi T, et al. High-resolution Schottky CdTe diode for hard X-ray and gamma-ray astronomy. Nucl Instrum Methods Phys Res Sect A: Accel Spectrom Detect Assoc Equip. 1999;436(1):111–9.
6. Patra JC, Kot AC. Nonlinear dynamic system identification using Chebyshev functional link artificial neural networks. Syst Man Cybern Part B: Cybern IEEE Trans. 2002;32(4):505–11.
7. Li S-J, Liu Y-X. An improved approach to nonlinear dynamical system identification using PID neural networks. Int J Nonlinear Sci Numer Simul. 2006;7(2):177–82.
8. Rafiee J, et al. Intelligent condition monitoring of a gearbox using artificial neural network. Mech Syst Signal Process. 2007;21(4):1746–54.
9. Chu JL, Li HY, Chen XJ. Research on improved bp learning algorithm of BP neural network. Adv Mater Res. 2013;765(4):1644–7.
10. Patra JC, et al. Identification of nonlinear dynamic systems using functional link artificial neural networks. Syst Man Cybern Part B: Cybern IEEE Trans. 1999;29(2):254–62.
11. Lin Y-Y, Chang J-Y, Lin C-T. Identification and prediction of dynamic systems using an interactively recurrent self-evolving fuzzy neural network. Neural Netw Learn Syst IEEE Trans. 2013;24(2):310–21.
12. Sicuranza GL, Carini A. On the BIBO stability condition of adaptive recursive FLANN filters with application to nonlinear active noise control. Audio Speech Lang Process IEEE Trans. 2012;20(1):234–45.
13 Lu H-j, Zhang H-m, Ma L-h. A new optimization algorithm based on chaos. J Zhejiang Univ sci A. 2006;7(4):539–42.
14. Bouras A, Syam WP. Hybrid chaos optimization and affine scaling search algorithm for solving linear programming problems. Appl Soft Comput. 2013;13(5):2703–10.

Part VII
Electronics Technology

Chapter 81
Analysis of the Magnetic Field Homogeneity of Two Different Geometric Configurations of the Helmholtz Coils by Simulation

Xiang Zhan, Guyin Rao, Jie Yuan and Yingying Li

Abstract In order to make a more direct and profound comparison of the magnetic fields generated in the central region of the circular Helmholtz coils and the square Helmholtz coils, simulation was conducted regarding the magnetic fields generated in these two situations respectively. The basic theory of magnetic fields generated by square and circular coils along their longitudinal axes is given in Sect. 1 of this chapter. Subsequently, the results of simulation adopting the finite element method (FEM) are compared with the theoretical value, which shows that the simulated result is well consistent with the theoretical data. A mean error of as low as 0.78 % in the range of $|y| < 20$ mm along the longitudinal axis of the circular Helmholtz coil is obtained while the corresponding result of square Helmholtz coil is merely under 1.03 %. Finally, it is concluded that the square Helmholtz coils can generate a larger uniform magnetic field.

Keywords Circular coil · Finite element analysis · Helmholtz coil · Magnetic field homogeneity · Square coil

81.1 Introduction

A uniform magnetic field is required in many instruments and systems. With its simple mechanical structure and capability of generating a large uniform magnetic field, the Helmholtz coil is widely used in many laboratory applications [1].

Numerical methods and analytical methods are two major approaches for the calculation of the electromagnetic field. However, since electromagnetic field can hardly be expressed by simple functions on many occasions, adding to the difficulty to obtain exact analytical solution, numerical methods have become the commonly used methods for calculating the distribution of electromagnetic field.

X. Zhan (✉) · G. Rao · J. Yuan · Y. Li
College of Optoelectronic Science and Engineering, National University of Defense Technology, 410073 Changsha, China
e-mail: zhanxiang09@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_81

The finite element method (FEM) is a widely used numerical method, which was initially proposed by Courant in 1943. The FEM works by firstly dividing the solution region into many small subregions which are usually referred to as “unit” or “finite element,” then using governing equations to calculate in each unit, next assembling all the results in each unit and obtaining the overall solution to the whole solution region eventually. With its high accuracy of calculation and being capable of applying to problems involving irregularly shaped boundaries, it has become an effective tool for engineering analysis [2–4].

Considering the difficulty of expressing the sophisticated magnetic flux density in the Helmholtz coils with elementary functions, in this chapter, to attain a better understanding of the magnetic fields generated by these two kinds of coils at the central zone, ANSYS Maxwell was utilized to analyze the magnetic fields generated respectively in same-sized circular and square Helmholtz coils. Simulation results were compared with the theoretical value, and uniform magnetic fields generated at the center by two kinds of coils during the simulation were also compared with each other.

81.2 Configurations of Coils

It is assumed that the radius and the breakover current of one single circular coil are R and I, respectively (Fig. 81.1).

In Fig. 81.1, y is the distance between one point on the central axis and the center O of the coil.

According to Bio–Savart law [5], the expression of the magnetic field at any point on the coil’s central axis is given by:

$$B(y) = \frac{\mu_0 I R^2}{2\sqrt{\left(y^2 + R^2\right)^3}} \tag{81.1}$$

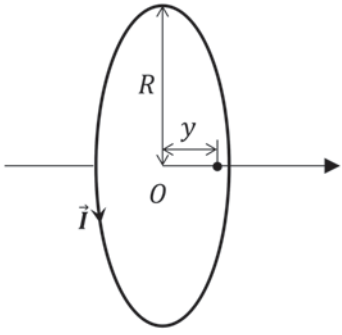


Fig. 81.1 The model of one single circular coil

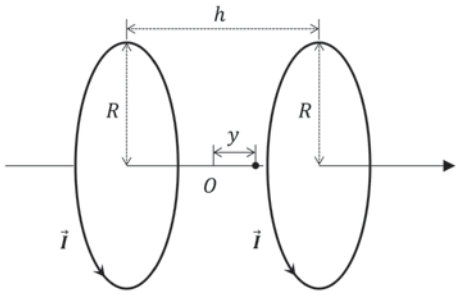


Fig. 81.2 The model of a pair of circular coils

Considering the circular coil, a pair of two identical, paralleled, coaxial coils whose have the separation distance of h are shown in Fig. 81.2.

The expression of the magnetic field at any point on their central axis is [6]:

$$B(y)=\frac{\mu_0 IR^2}{2}\left\{\left[\left(\frac{h}{2}-y\right)^2+R^2\right]^{-3/2}+\left[\left(\frac{h}{2}+y\right)^2+R^2\right]^{-3/2}\right\}. \tag{81.2}$$

The expression Eq. 81.2 can be expressed by a Taylor series near the center:

$$B(y)=B(0)+\frac{1}{2}B^{(2)}(0)y^2+\frac{1}{24}B^{(4)}(0)y^4+\cdots, \tag{81.3}$$

where $$B(0)=\frac{8\mu_0 I}{5\sqrt{5}R},\ B^{(2)}(0)=3\mu_0 IR^2\left(\frac{h^2-R^2}{\left(R^2+h^2/4\right)^{7/2}}\right).$$

When the separation distance h is equal to the radius R, which is the feature of circular Helmholtz coil, the second derivative of the Taylor series is equal to zero, therefore a widest uniform magnetic field can be obtained [7].

Considering the square coil whose side length is 2 l and separation distance is a (Fig. 81.3).

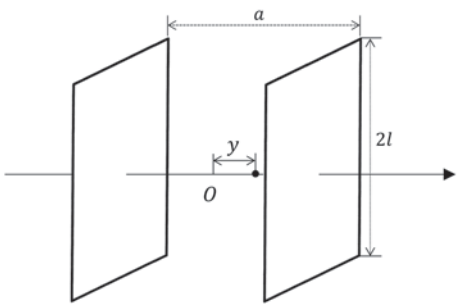


Fig. 81.3 The model of a pair of square coils

Then the expression of the magnetic fields at any point of its symmetry axis is given by:

$$B(y)=\frac{2\mu_0 Il^2}{\pi}\left(\frac{1}{[l^2+(y+a^2)]\sqrt{2l^2+(y+a)^2}}+\frac{1}{[l^2+(y-a^2)]\sqrt{2l^2+(y-a)^2}}\right) \tag{81.4}$$

We can also expand the expression in a Taylor series around $y=0$, and when the separation distance a of the two square coils is 0.5445 l, the second derivative of the Taylor series is equal to zero, then the condition of the square Helmholtz coils is fulfilled [8].

In order to identify the different homogeneity of the magnetic field, the expression as follows is used:

$$H=\frac{B-B(0)}{B(0)}*100\% \tag{81.5}$$

where B is the magnetic field intensity at any point in the coils and $B(0)$ is the magnetic field intensity at the center of the coils.

Normally it is expressed as a percent difference (%), from which the homogeneity of magnetic field can be measured.

81.3 Simulation

ANSYS Maxwell, based on Maxwell differential equations, is a widely used commercial software, which can convert the calculation of electromagnetic fields in engineering problems to the solution of matrixes of low-frequency electromagnetic fields, through the finite element analysis method and the automatic adaptive meshing techniques. To perform the analysis of the magnetic fields of these two coil pairs, their geometric sizes are set up at the same value, the radius of circular coil pair is 50 mm and the side length of square coil pair is 100 mm (Fig. 81.4).

81.3.1 Comparison Between the Simulated and Calculated Data

The magnetic field distribution on the Y-axis ($|y|$ Equ < Equ 25 mm) of circular and square coils according to the simulated data is shown in Fig. 81.5, from which significant area of uniform magnetic field in both the two coils can be obtained.

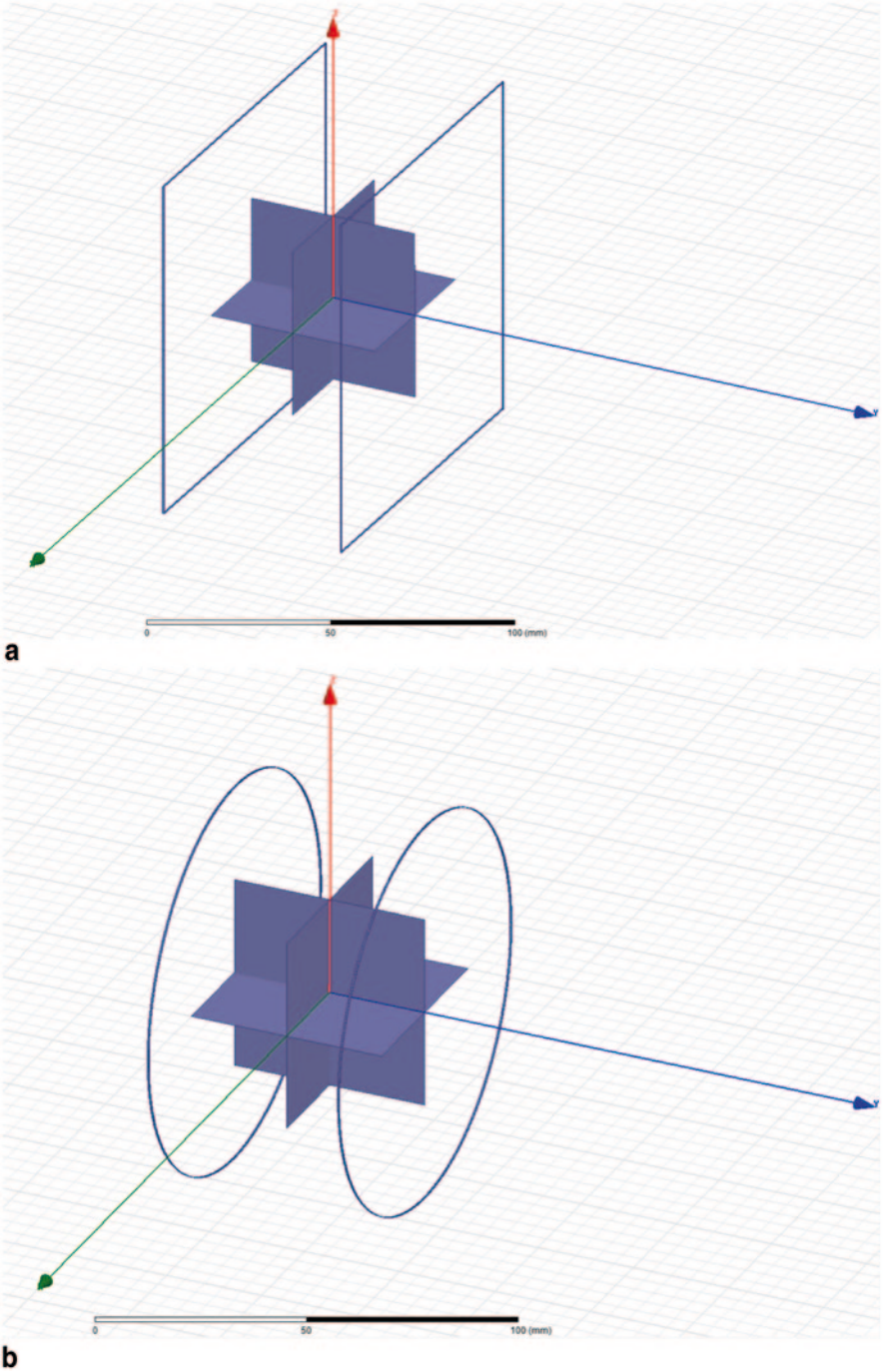


Fig. 81.4 **a** The calculation model of square coils pair. **b** The calculation model of circular coils pair

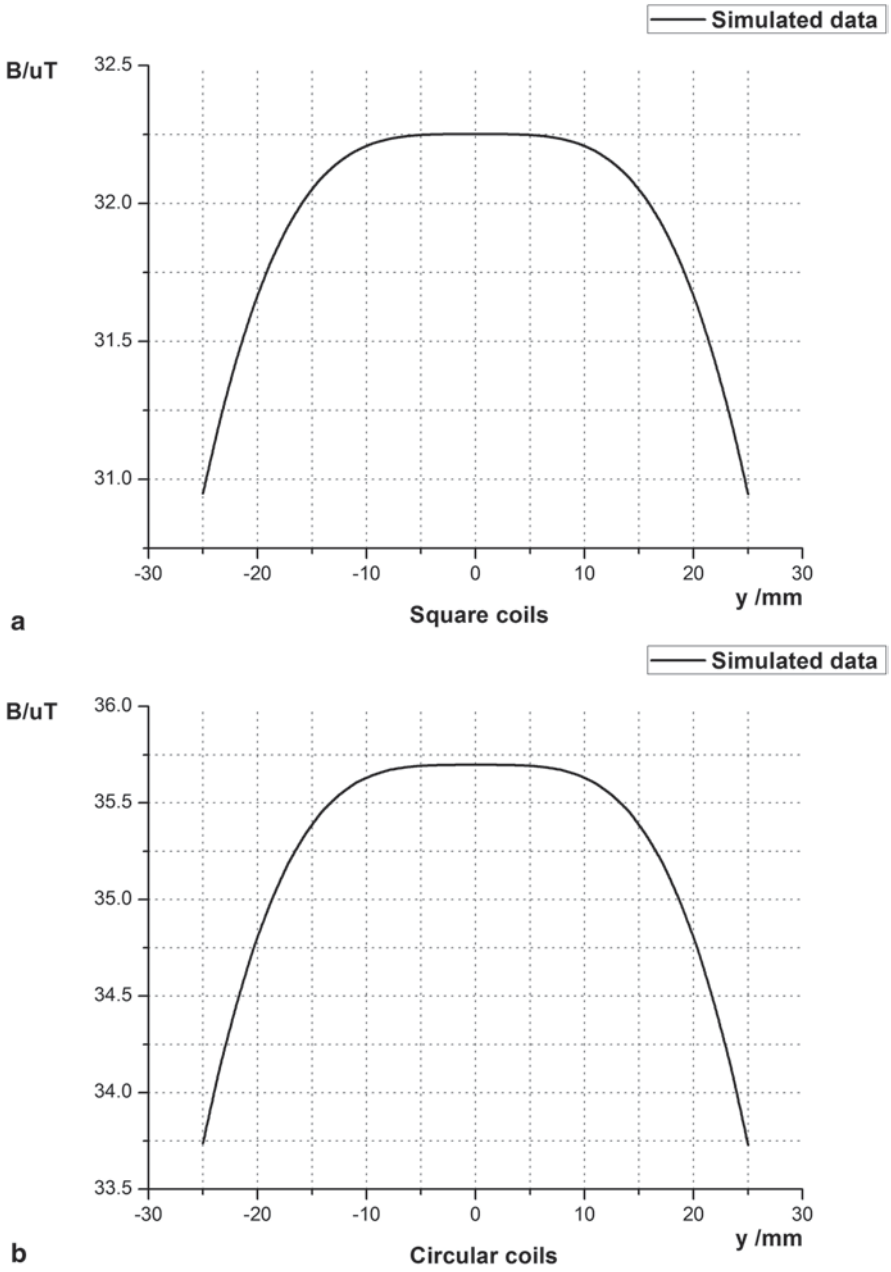


Fig. 81.5 The magnetic field distribution on the *Y*-axis of square and circular coils. **a** Square coils. **b** Circular coils

Table 81.1 Comparison between the simulated and calculated data

Circular Helmholtz coils				Square Helmholtz coils		
|y|/mm	*Calculated data/uT*	*Simulated data/uT*	*Error/%*	*Calculated data/uT*	*Simulated data/uT*	*Error/%*
0	35.97	35.70	0.751	32.57	32.25	0.993
5	35.96	35.69	0.753	32.57	32.25	0.995
10	35.90	35.63	0.763	32.53	32.21	1.002
15	35.66	35.39	0.777	32.38	32.05	1.013
20	35.08	34.80	0.790	31.99	31.67	1.031

With Eqs.81.2 and 81.4, we can get the calculated magnetic field distribution on the *Y*-axis easily. And the result of the comparison of the simulated data and calculated data is listed in Table 81.1.

From Table 81.1, we can find that the error is proportional to the distance referred to the center of separation (y). Noteworthy is that the maximum error in the simulated data of square coils is only 1.03 % ($y=20$). Comparatively, the error of the normalized data is under 0.04 %. Therefore, these results can reflect and demonstrate the magnetic fields distribution well.

81.3.2 Comparison of the Uniform Magnetic Fields in Circular and Square Coils

Both of the two coils can generate uniform magnetic fields in theirs centers; however, the differences in the geometry of the two coils may contribute to the different areas of the uniform magnetic field (Fig. 81.6).

Figure 81.6 illustrates the magnetic field distribution on XOY plane, from which it can be noticed that more intensive outlines of magnetic field levels in the circular coils are obtained.

Aiming to compare the uniform magnetic fields generated in two kinds of coils, evaluation lines are drawn respectively on *X*-axis and *Y*-axis, as well as the diagonal line, in the simulation. The distribution of magnetic fields on these evaluation lines is shown in Fig. 81.7.

Table 81.2 depicts the homogeneity of the magnetic field of the square and circular coils. From the simulation, it can be found that better homogeneity of magnetic field can be achieved in the square Helmholtz coils particularly on *Y*-axis and diagonal line.

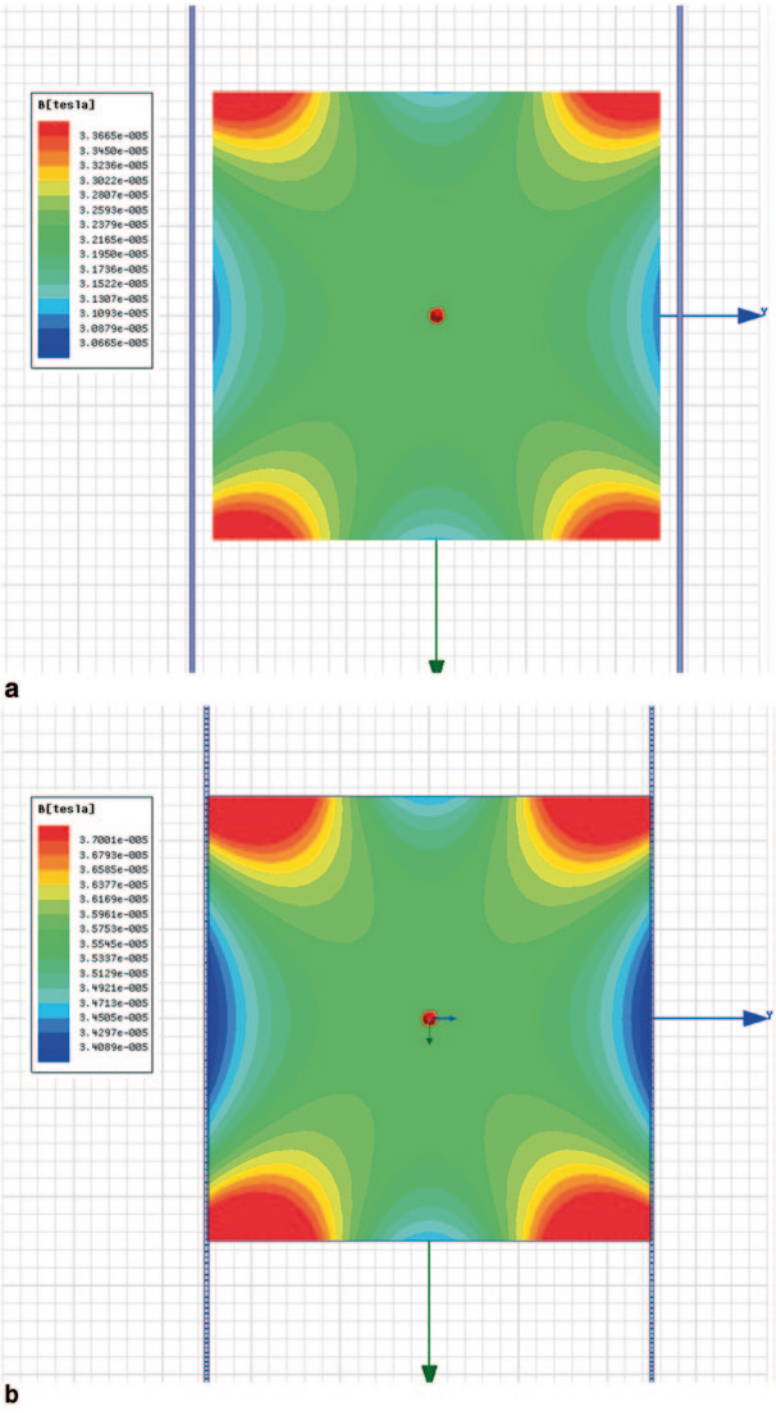


Fig. 81.6 The magnetic field distribution on XOY plane of square and circular coils. **a** Square coils. **b** Circular coils

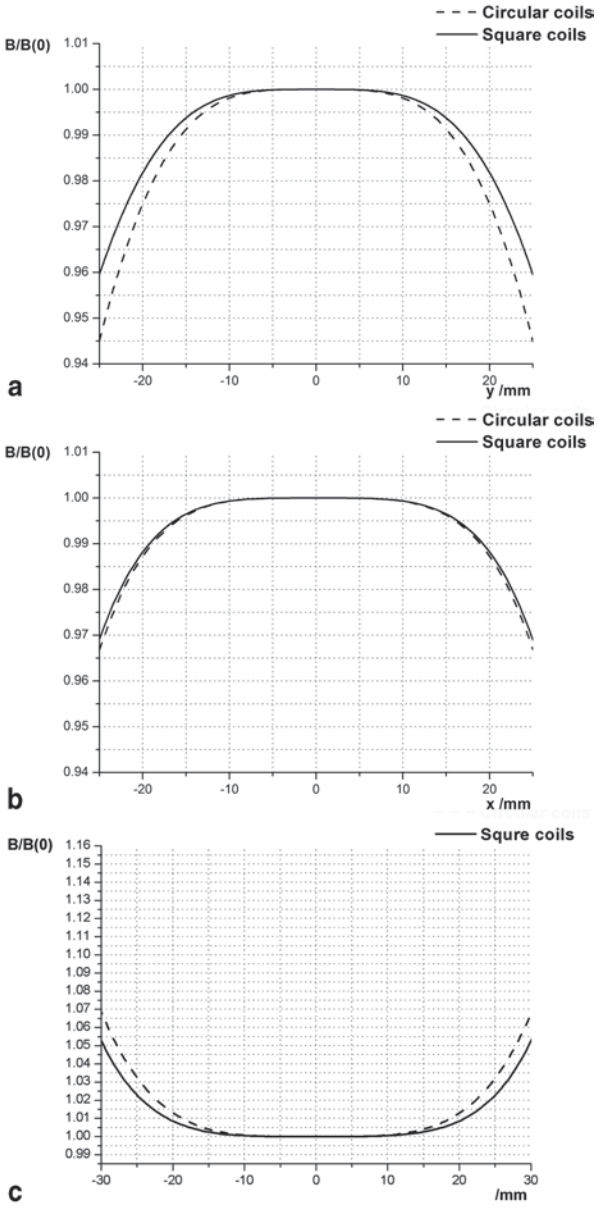


Fig. 81.7 The distribution of magnetic fields (normalized) on the evaluation lines. **a** *Y*-axis. **b** *X*-axis. **c** Diagonal line (on XOY plane)

Table 81.2 The homogeneity of the magnetic field of the square and circular coils

	Range of uniform magnetic field of circular coils			Range of uniform magnetic field of square coils		
Homogeneity/%	*Y-axis/mm*	*X-axis/mm*	*Diagonal line/mm*	*Y-axis/mm*	*X-axis/mm*	*Diagonal line/mm*
0.1	16.99	21.90	21.41	19.02	22.68	24.43
0.5	25.90	32.13	31.70	28.95	33.33	35.29
1	31.18	37.88	37.43	34.06	38.47	41.46
1.5	34.79	41.59	41.45	37.97	42.30	45.53

81.4 Conclusion

FEM is a feasible and effective way to explore some sophisticated magnetic flux density and magnetic distribution. According to the simulation performed in this chapter, it can be concluded that the square Helmholtz coils produce better homogeneity of magnetic field compared with the circular Helmholtz coils with the same geometry size. Furthermore, both the differences between square and circular coils and the relationship between the homogeneity of magnetic fields and the physical parameters of the coils should be taken into consideration in practical applications.

References

1. Song Q, Wang W, Zhang J. Study on the uniform magnetic field in Helmholtz coils by simulation. Saf EMC. 2010;10(2):61–4 (In Chinese).
2. Malik B. Electric field calculations by numerical techniques. Rourkela: National Institute of Technology Rourkela. 2009.
3. Boughrara K, Ibtiouen R, Touhami O, Zarko D. Analytical and numerical calculation of magnetic field distribution in the slotted air-gap of tangential surface permanent-magnet motors. Serb J Electr Eng. 2009;6(1):15–31.
4. Wimmer G, Clemens M, Lang J. Calculation of magnetic fields with finite elements. In: Breitner MH, Denk G, Rentrop P, editors. From Nano to Space. Berlin: Springer; 2008. pp. 111–24.
5. Bell GB, Marino AA. Exposure system for production of uniform magnetic fields. Electromagn Biol Med. 1989;8(2):147–58.
6. Cvetkovic D, Cosic I. Modelling and design of extremely low frequency uniform magnetic field exposure apparatus for in vivo bioelectromagnetic studies. In IEEE EMBS 2007, IEEE. 2007. pp. 1675–8.
7. Wang J, She S, Zhang S. An improved Helmholtz Coil and analysis of its magnetic field homogeneity. Rev Sci Instrum. 2002;73(5):2175–79.
8. Alvarez AFR, Franco-Mejia Edinson, Pinedo-Jaramillo CR. Study and analysis of magnetic field homogeneity of square and circular Helmholtz coil pairs: a taylor series approximation. In Andean Region International Conference (ANDESCON), 2012 VI, IEEE, 2012. pp. 77–80.

Chapter 82
A Multichannel High-Precision Pulse Counting System Based on FPGA

Yuan Li, Yunfei Zhou and Yitao Li

Abstract This chapter presents a multichannel and high-precision pulse counting system based on the time interval measurement on the Xilinx ZYNQ-7000 series device. The theoretical resolution of time interval measurement can reach 10-ps using the internal fast carry chain on the field-programmable gate array (FPGA) and the wave union method. As a result, as limited by the hardware condition and the carry chain's structure, the practical resolution of 40-ps is realized, which means that 2048 times subdivision for pulse signal running at 12M is available. It fully meets the requirement of pulse counting in the measurement system of the precise laser interferometer position.

Keywords Pulse counting · Time interval · FPGA · Fast carry chain · Wave union

82.1 Introduction

In the field of ultraprecision position measurement, the laser heterodyne interferometry method is widely used due to many advantages such as fast response, large measuring range, high signal-to-noise ratio, and insensitivity to the change of light intensity. Electro-optic, acousto-optic, machinery, and other technical realization of optical frequency modulation constitute a variety of different laser interference measurement systems. The laser heterodyne interferometry system based on the Zeeman effect uses the Zeeman double-frequency laser device to issue two beams of linear

Y. Li (✉) · Y. Zhou
School of Mechanical Science and Engineering, Huazhong University of Science and Technology, 430074 Wuhan, China
e-mail: ly1991@hust.edu.cn

Y. Li
Hangzhou Dahe Thermo-magnetics Co. Ltd., 310053 Hangzhou, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_82

polarized light with different frequencies (f1, f2) and polarization directions. On the one hand, they form the reference light signal with the frequency of f1−f2 resulting from the spectroscope inside laser device; on the other hand, the linear polarized light with the frequency of f1 is obtained by reflecting from the fixed reference plane mirror, and the linear polarized light with the frequency of f2 ± Δf is obtained by reflecting from the measuring plane mirror installed on the measured object (Δf refers to the additional frequency when the measurement mirror is moving because of the Doppler effect, and plus or minus means the direction of the movement). Both form the measured light signal with the frequency of f1−f2 ± Δf, therefore, the displacement of the measured object is transformed to the frequency difference between reference light signal and measured light signal as shown in formula 1:

$$S = \int_0^\tau v dt = \int_0^\tau \frac{\Delta f \times c}{4f2} dt = \frac{\lambda}{4} \int_0^\tau \Delta f dt \tag{82.1}$$

$$\Delta f = \frac{\Delta N_m - \Delta N_f}{t}, \tag{82.2}$$

where v is the velocity of measured mirror, c is the velocity of laser, λ is the laser wavelength, S is the displacement of measured objects, ΔN_m is the increment of the cycle count of the measured signal, ΔN_f is the increment of the cycle count of the reference signal, and t is the sample interval.

The reference light signal and measured light signal are received by the laser receiver and then transformed into the square wave signals by the photoelectric conversion and corresponding modulation. To calculate Δf, as shown in formula 2, it needs to sample the ΔN_m and the ΔN_f at a certain time interval t.

The resolution of cycle number is critical to the requirements of high precision. That is to say, high-magnification subdivision technology is the key technology of the laser heterodyne interferometry system. To realize the position measurement precision of nanometer level, 2048 times subdivision is needed. Taking interference light signal at 12M, for example, it needs to counter at up to 24 GHz, which is impossible on hardware at present; in this sense, it is necessary to research the effective subdivision method.

Several decades ago, the main methods used to implement the time interval measurement with the picosecond resolution were based either on the analog time stretching principle or on the time-to-amplitude conversion. For example, E. Raisanen-Ruotsalainen presented a time interval measurement method with an integrated complementary metal–oxide–semiconductor (CMOS) circuit, which combined a time-to-voltage converter (TVC) and a dual-slope analog-to-digital (A/D) converter and obtained the resolution of 200 ps or better [1]. P. Chen proposed a low-cost and low-power CMOS time-to-digital converter (TDC) with 50-ps resolution by both methods combined with each other [2]; however, these methods always obtain high resolution at the expense of long conversion time, significant circuit complexity, and limited stability. At present, rapid progress in terms of integrated circuit has promoted the development of fully integrated TDC based on the digital

delay lines. In 2000, the resolution of 100-ps was obtained because of the new design of the field-programmable gate array (FPGA) delay lines by Szplet et al. [3]. In 2009, the resolution of 55-ps on intervals less than 20-μs was obtained by Aloisio et al. [4]. In 2010, a TDC with 24-ps resolution was implemented on Virtex-II Pro FPGA by Daigneault et al. [5].

In conclusion, TDCs' implementation on FPGA has proven to be efficient and successful to achieve high resolution in the time interval measurement system, which could significantly improve the performance of electronics subdivisions; therefore, this method is adopted in this chapter.

82.2 Implementation Method

The cycle counting of the reference signal and measured signal is made up of two parts: the coarse counting and the fine counting. Taking the measured signal, for example, the coarse counting aims at the integer part of ΔN_m, while the fine counting aims at the decimal part.

82.2.1 Coarse Counting

In the coarse counting part, an ordinary binary counter is selected to cooperate with the related buffer so as to complete the coarse counting because it has a simple structure and is less resource-intensive. The rising edge of the measured signal (Meas) is taken as trigger condition of the counter, and the rising edge of the sample signal (Samp) is taken as that of the buffer. But the carry delay from low bit to high bit leads different bits to flip asynchronously so that the counter will be unreliable in a short time interval upon the rising edge of the measured signal, and the unreliable state is called metastable state. If the sample signal hits the metastable interval, the buffer will store an uncertain value, the so-called carry error.

In this chapter, it begins to check the status of the measured signal on the rising edge of the sample signal. As shown in Fig. 82.1, once the status is 1, the result of

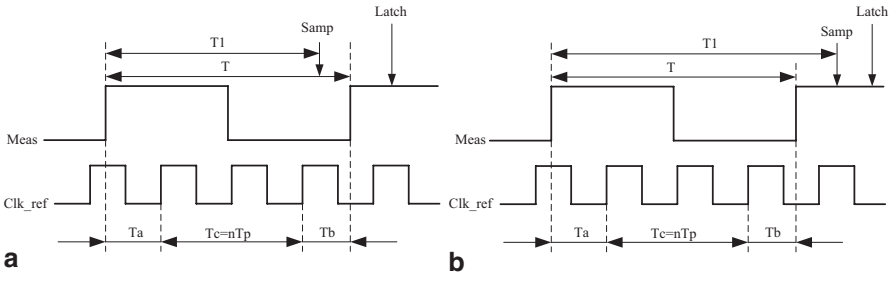


Fig. 82.1 Two kinds of cases in the coarse counting part and a measurement method of the time interval *T*

coarse counting will be stored after a certain time, which is longer than the metastable interval. As shown in Fig. 82.1a, if the rise edge of the sample signal hits the measured signal's low level interval, the actual result of the coarse counting N refers to the result of buffer (Q) minus 1, while the actual result of the fine counting is $T1/T$. If not, the actual result of the coarse counting N is Q, while the actual result of the fine counting is $T1/T$ minus 1, as shown in Fig. 82.1b. Above all, it can be concluded that the count N of measured signal $N = Q - 1 + T1/T$.

82.2.2 Fine Counting

In the fine counting part, it adopts the high frequency clock (Clk_ref) running at 500 MHz as the standard to implement the measurement of the time interval T and $T1$. Taking the time interval T, for example, as shown in Fig. 82.1, it consists of a synchronous part Tc and asynchronous parts Ta and Tb. For the same reason, the ordinary binary counter is selected for the cycle number of the reference clock and the delay from internal carry chain on FPGA is used for the precise measurement of Ta and Tb, which forms a wide-range and high-resolution time interval measurement. The high-speed counter is responsible for the cycle counting of the reference clock, and the result in cycle buffer 2 is transferred to cycle buffer 1. The current counting result is stored in cycle buffer 2 on the rising edge of the measured signal. So the reference clock cycle number n can be calculated by above two values, $Tc = nTp$. Meanwhile, the reference clock signal was fed into the fast carry chain and the result in carry buffer 2 is transferred to carry buffer 1, and carry status is stored in carry buffer 2 on the rise of the measured signal, which respectively are the thermometer code about Ta and Tb. Above all, it is concluded that $T = Tc + Ta + Tb$. In the same way, the time interval $T1$ is available.

For the measurement of Tc, this chapter uses two high-speed counter (Count_p and Count_n) and two buffers (Cnt_p and Cnt_n) to avoid the carry error. The rising edge of the reference clock triggers the Count_p counter, while the falling edge triggers the Count_n counter. And the result of Count_p is stored in Cnt_p on the rising edge of the reference clock, while the result of Count_n is stored in Cnt_n on the falling edge. As shown in Fig. 82.2, Cnt_n is always reliable in the high-level interval of the reference clock signal, and Cnt_p is always reliable in the low-level interval of the reference clock signal. According to the instantaneous status of the reference clock signal on the rising edge, reliable counting result is selected as the Qs; however, it is effective only depending on the correct instantaneous status of the reference clock signal. When the rising edge of the measured signal hits close to the edge of the reference clock signal, the instantaneous status of the reference clock signal may be wrong.

In order to solve this problem, as shown in Fig. 82.2a, the first output bit (Carry_1) of the tapped delay line is designed to delay t relative to the reference clock signal, which is longer than the setup and hold time of the trigger to ensure that Carry_1 can be latched correctly when incorrect instantaneous status happens. It assumes

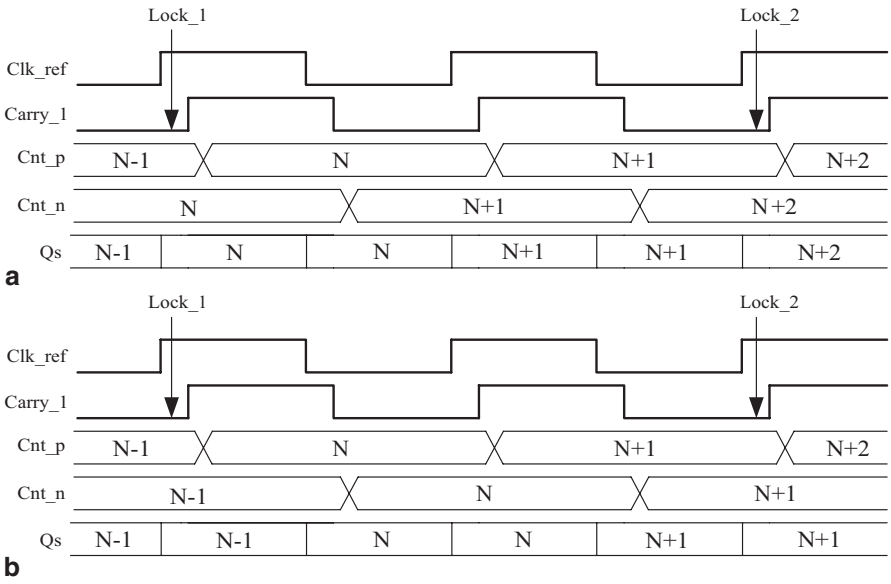


Fig. 82.2 Two kinds of cases of the high-speed counters for the measurement of *Tc*

that both the edges of the measured time interval are Lock_1 and Lock_2. When Lock_1 hits the position where the status of Clk_ref and Carry_1 are 1 and 0, the calculation result of *Ta* + *Tb* will be smaller than the expected result of one reference clock cycle, while when Lock_2 hits the position, it will lead to the calculation result of *Ta* + *Tb,* which is larger than the expected result of the reference clock cycle. But it can only offset the count error of *Qs* under the condition as shown in Fig. 82.2(1). Therefore, the counter reset is essential to ensure that the Count_n counter starts to work before Count_p.

For the measurement of *Ta* and *Tb,* it is realized by the tapped delay line consisting of the dedicated fast carry chain on the Xilinx ZYNQ-7000 series XC7Z-020CLG484 device. In every configurable logic block (CLB), there are two independent carry chains having a height of four bits and is running upward and aiming at performing fast arithmetic addition and subtraction with the carry delay as low as 10 ps. In this chapter, it cascades 28 carry chains in adjacent CLBs, but the cascading of any two CLBs will bring extra delay up to 35 ps, called the ultrawide. Wu et al. put forward and implemented the wave union methods to improve the resolution by tapped delay line [6–7], a ring oscillator is released on the rise edge of the sample signal, and on the next 16 clock rising edge. The output of the carry chain will be latched and analyzed to realize measurement resolution of 10 ps. As shown in Fig. 82.3, Carry4 is the independent carry chain on the XC7Z020CLG484. The CI input and COUT output refer to the lowest carry bit and the highest carry bit of carry chain unit, respectively, which are used to cascade the adjacent carry chain. The DI input is one operand of addition or subtraction. The S inputs are the XOR

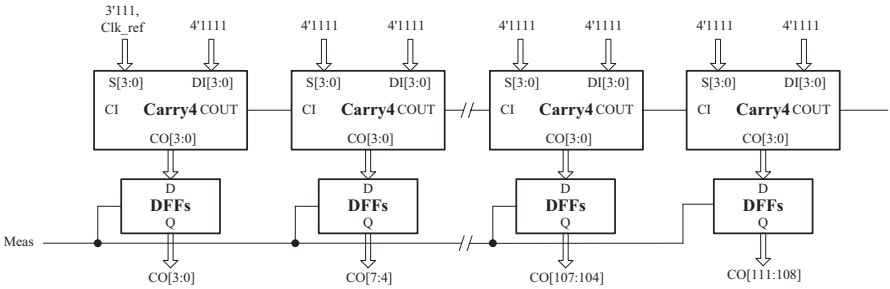


Fig. 82.3 Tapped delay line architecture composed of the fast carry chains

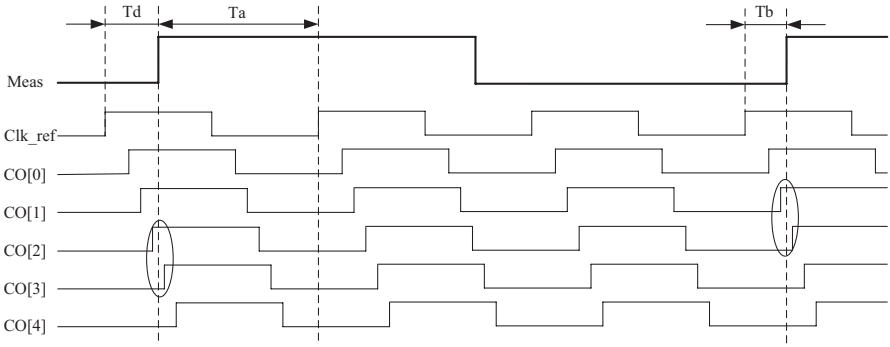


Fig. 82.4 Waveforms of the propagate signals from CO outputs of the carry chains

of both operands. The CO outputs are carry outputs of the carry chain. To propagate the reference signal, the first two operands are 1 and the reference signal value and others are 1 and 0, respectively. Figure 82.4 shows that the waveforms of CO outputs would be generated. By finding 1–0 logic transition on the rising edge of the measured signal in combination with the carry delay value, *Td* will be figured out. $Ta = Tp - Td$. *Tb* will be figured out for the same reason.

82.3 Results

From what was mentioned above, the precision of the pulse counting system in this chapter is determined by the time interval measurement method; therefore, both the results of the time interval measurement and the pulse count are presented. In the experimental part, the laser counting board by independent research and development is chosen as the hardware platform. The high-precision clock chip (CDCM61001) is chosen to produce the reference clock. The internal integrated

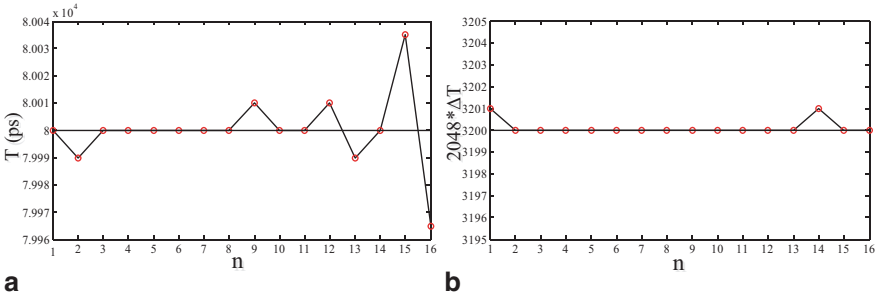


Fig. 82.5 Experimental results of the time interval *T* and pulse count increment

high-precision phase locked loop (PLL) on FPGA is used to produce a clock signal running at 12.5M and another clock signal running at 8M as the measured signal (Meas) and the sample signal (Samp), respectively. The Xilinx integrated with logic analysis tool ChipScope Pro is adopted to meet the requirement of multi-node logic analysis. Upon experimenting for many times, a part of the results of the time interval *T* (80000-ps) and general count increment 2048*Δ*N* (1.5625) are presented in Fig. 82.5. Figure 82.5a shows 16 groups of experimental results about *T,* and Fig. 82.5b shows the correspondent results of 2048*Δ*N*. From the two figures, it can be concluded that the biggest random measurement error of the time interval is only 35-ps and 14-ps (RMS), which may result in the pulse count error of 1/2048.

The factors which affect the measurement precision and its correspondent solutions are presented below:

1. Stability and jitter of the reference clock cycle

The jitter of the reference clock directly affects the measurement precision for the part of *Ta* and *Tb*. In this chapter, the higher precision clock chip (CDCM61001) is chosen, whose general jitter meets 1-ps(RMS), 27-ps(p-p).

2. Delay from internal routing

In the fine measurement part, the small route delay exists for different CLBs, although the sample signal enters the global clock network. It may greatly reduce the route delay. Thus, the tapped delay line is designed as short as possible. Meanwhile, as to the measurement of *Tc,* the Count_p and Count_n are placed strictly symmetrical to reduce the route distance as soon as possible.

3. Stability of carry delay

The carry delay is related to the temperature and the operation voltage. Using the stable power supply or the effective filter will be greatly helpful. In this way, heat dissipation needs attention..

82.4 Conclusion

In this chapter, a high-accuracy multichannel pulse counting system is designed and realized based on the time interval measurement with the theoretical precision of 40-ps using the internal carry chain on XC7Z020CLG484. In fact, the experimental resolution of the time interval measurement can satisfy the high-resolution requirements of pulse counting in the laser interference precise position measurement system. Meanwhile, it is necessary to implement interpolation of the ultrawide to further improve the interval resolution. According to Wu et al., it needs a very accurate internal oscillation, and it consumes more logical resources. In order to overcome these shortcomings, two or more external reference clock signals with small fixed phase differences can be designed to feed into the carry chain; therefore, how to produce the reference clock signals with low jitter and small fixed phase differences becomes the main content in the further research.

References

1. Raisanen-Ruotsalainen E, Rahkonen T, Kostamovaara J. Time interval measurements using time-to-voltage conversion with built-in dual-slope A/D conversion. IEEE Int Symp Cir Syst. 1991;5:2573–6.
2. Chen P, Chen G, Shen Y. A low-cost low-power CMOS time-to-digital converter based on pulse stretching. IEEE Trans Nucl Sci. 2006;53:2215–20.
3. Szplet R, Kalisz J, Szymanowski R. Interpolating time counter with 100 ps resolution on a single FPGA device. IEEE Trans Instrum Meas. 2000;49(4):879–83.
4. Aloisio A, Branchini P, Giordano R, Izzo V, Loffredo S. High-precision time-to-digital converter in a FPGA device. IEEE Nuclear Science Symposium Conference Record, 2009 IEEE. IEEE; 2009. p. 290–4.
5. Daigneault M, David J. A novel 10 ps resolution TDC architecture implemented in a 130 nm process FPGA. NEWCAS Conference, 2010 8th IEEE International, IEEE; 2010. p. 281–4.
6. Bayer E, Zipf P, Traxler M. A multichannel high-resolution (< 5 ps RMS between two channels) time-to-digital converter(TDC) implemented in a field programmable gate array(FPGA). IEEE Nuclear Science Symposium Conference Record, 2011 IEEE, IEEE; 2011. p. 876–9.
7. Wu J, Shi Z. The 10-ps wave union TDC: improving FPGA TDC resolution beyond its cell delay. IEEE Nuclear Science Symposium. Conference Record, 2008 IEEE, IEEE. Oct. 19–25; 2008. p. 3440–6.

Chapter 83
High-Precision Ultrasonic Flow Measurement Based on GP22

Ling Kang, Hui Lin, Tianyu Zhang and Zhao Deng

Abstract With the extensive utilization of ultrasonic flow meters in industry and commerce, the demand of improving accuracy of measurement has become increasingly urgent. This chapter presents a new measuring tubing that is a combination of circular tube and square tube. According to the results of computational fluid dynamics (CFD) simulation and experiment, this proposed measuring tubing can produce a more even velocity distribution of liquid than the traditional circular tubing. Accompanied by the tubing, a novel high-precision ultrasonic flow-measuring circuit based on the time-digital converter (TDC) of GP22 is designed. This measuring circuit, which depends on the principle of time difference method, can not only obtain accurate measuring data of instantaneous flow after calibration but also reduce power consumption.

Keywords Ultrasonic · Flow · Measurement · Tubing · Circuit · TDC-GP22

83.1 Introduction

In the past, the mechanical flow meter and the electromagnetic flow meter occupied the majority of the domestic market; however, both of them had significant drawbacks, such as short life span, requirement of high-quality liquid, and so on [1], which thus enabled further improvement in the level of ultrasonic flow measurement.

In recent years, the ultrasonic flow meter has played a more and more important role in the field of flow-measuring instruments. Compared with other flow meters, its characteristics can meet market demands, such as stability, long life span, etc. Most importantly, it is easier to achieve high-precision results.

H. Lin (✉) · L. Kang · T. Zhang · Z. Deng
Department of Aeronautics, School of Physics and Mechanical & Electrical Engineering, Xiamen University, 361000 Xiamen, China
e-mail: 709535799@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_83

In this chapter, a new measuring tubing is proposed to achieve better accuracy and a high-precision chip of TDC-GP22 is adopted in the design. The design details are organized as below.

83.2 Basic Principles

Several methods can be applied in the ultrasonic flow meter, such as the noise method and the Doppler method while the method of time difference is mostly used.

Figure 83.1 illustrates the typical measuring principle of time difference method, the arrow v represents the fluid flow velocity and direction. The transducers TRA and TRB, positioned upstream and downstream, respectively can both emit and receive ultrasonic signals; and L represents the distance between them. In addition, the angle between the propagation path and velocity axial direction is represented by θ, and D stands for the diameter of pipe.

Therefore, the fluid velocity v can be calculated by the equation below [2]:

$$v \approx \frac{c^2 \Delta t}{2L\cos\theta}, \tag{83.1}$$

where Δt represents the transit time difference and c stands for ultrasonic velocity in fluid. Then, the instantaneous volume flow q can be expressed as below [3]:

$$q = \frac{\pi}{4} D^2 v. \tag{83.2}$$

Platinum resistance temperature sensor is used to measure the temperature of liquid. As a result, upon the measurement of the platinum resistance value R_t, the relevant temperature T can be obtained by the following equation:

$$R_t = R_0 * (1 + AT + BT^2), \tag{83.3}$$

where R_0 is the nominal resistance value when the temperature is 0 °C, and $A = 3.9083 \times 10^{-3}\,°C^{-1}$, $B = -5.775 \times 10^{-7}\,°C^{-2}$.

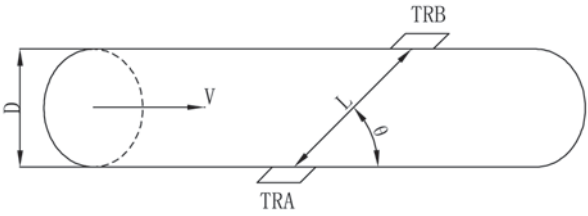


Fig. 83.1 Principle of time difference

83.3 Design and Simulation of Measuring Tubing

83.3.1 Structure Design

Regarding the ultrasonic flow measurement, the quality of tubing design has a great effect on precision and stability of measurement. The traditional circular tubing always needs at least one reflector for ultrasonic reflection. This structure may easily cause backflow and turbulence, thus causing low precision and unsteadiness.

To solve this problem, the new V-reflection tubing, which combines the circular tube and square tube, is proposed. As shown in Fig. 83.2, the inlet and outlet are both designed into circular tubes, and the section of the middle tube is designed into a square with circular arcs. Furthermore, the mounting holes of upstream and downstream ultrasonic transducers are designed in the middle tube. The angle between each axis of mounting holes and tubing axis is 30 °C, and the two mounting holes are designed in a symmetrical manner.

For any kind of the measuring tubing, the existence of a reflector will affect the fluency of liquid flow more or less. The greatest advantage of this new tubing is that the inside wall of square tube can be used for ultrasonic reflection instead of additional reflector. As nothing can block the flow of fluid in this tubing, the possibility of the backflow can be decreased to a minimum. Another advantage is that the residual quantity of impurities in this tubing can be minimized; besides, the phenomenon of turbulence can also be reduced as greatly benefited by the structure without reflector.

83.3.2 Fluid Simulation

The CFD numerical simulation software, Fluent 14.0, is used to simulate the velocity distribution of liquid flow in tubing. In addition to the new V-reflection tubing, the simulation also includes two kinds of traditional circular tubings for the purpose of comparison. The two circular tubings are the V-reflection circular tubing with one reflector and the U-reflection circular tubing with two reflectors.

The velocity profile of three tubings is shown in Fig. 83.3 with the inner liquid set as water, the entrance velocity of 2 m/s and the velocity direction from left to right [4].

Seen from the simulation result, only in the left profile, there is no backflow and turbulence in the propagation path of ultrasonic. Hence, the new V-reflection tubing

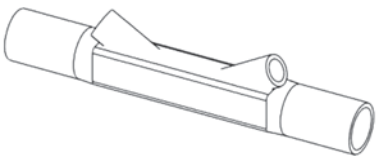


Fig. 83.2 Structure of the new tubing

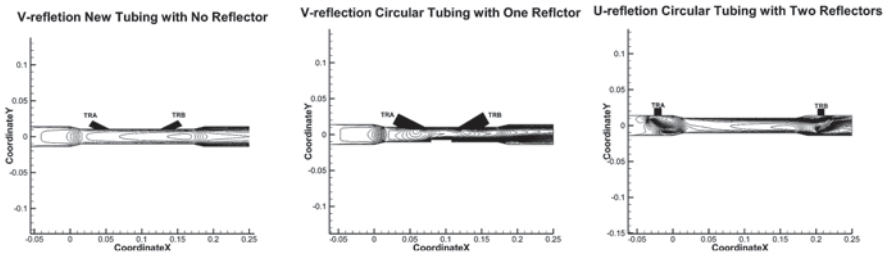


Fig. 83.3 Velocity profile of three tubings

without reflector can obtain a more even and stable velocity distribution than other two circular tubings.

83.4 Measuring Circuit Design

83.4.1 Hardware Design

As the core component of the ultrasonic flow meter, all other components need the support of a high-quality measuring circuit to realize their functions. The measuring circuit proposed by this chapter mainly consists of five parts: TDC-GP22, micro control unit (MCU), liquid crystal display (LCD), ultrasonic transducers, and platinum resistance temperature sensor. TDC-GP22 is the new generation of TDC made by ACAM. Its new function of first-wave detection makes it adapt the system to long-term signal attenuation and detect bubbles. This chapter adopts MC9S08LL-64CLK, an 8-bit microcontroller made by Freescale, as the MCU of its measuring circuit.

The power of the whole measuring circuit can be supplied purely with a 3-V battery. The working current of MCU under low-power mode is about 7.3 μA and the current consumption of LCD is about 100 μA; besides, the typical current consumption of the complete flow and temperature measuring unit in TDC-GP22, including the ultrasonic transducers and the temperature sensor, is about 2.3 μA [5]. Upon measurement, the power consumption of the whole measuring circuit is no more than 0.57 mW. To sum up, this measuring circuit really features low power.

As shown in Figs. 83.4 and 83.5, powerful internal integrated measuring and calculating components both simplify the peripheral circuits of MCU and TDC-GP22 [6]. MCU sends instructions regularly to TDC-GP22 with four-wire serial peripheral interface (SPI), and then TDC-GP22 generates signal pulses of 1 MHz frequency to the upstream ultrasonic transducer. As soon as the downstream ultrasonic transducer receives the signal pulses after propagation, the capability of first-wave detection gives MCU an indication of the strength of the received signals

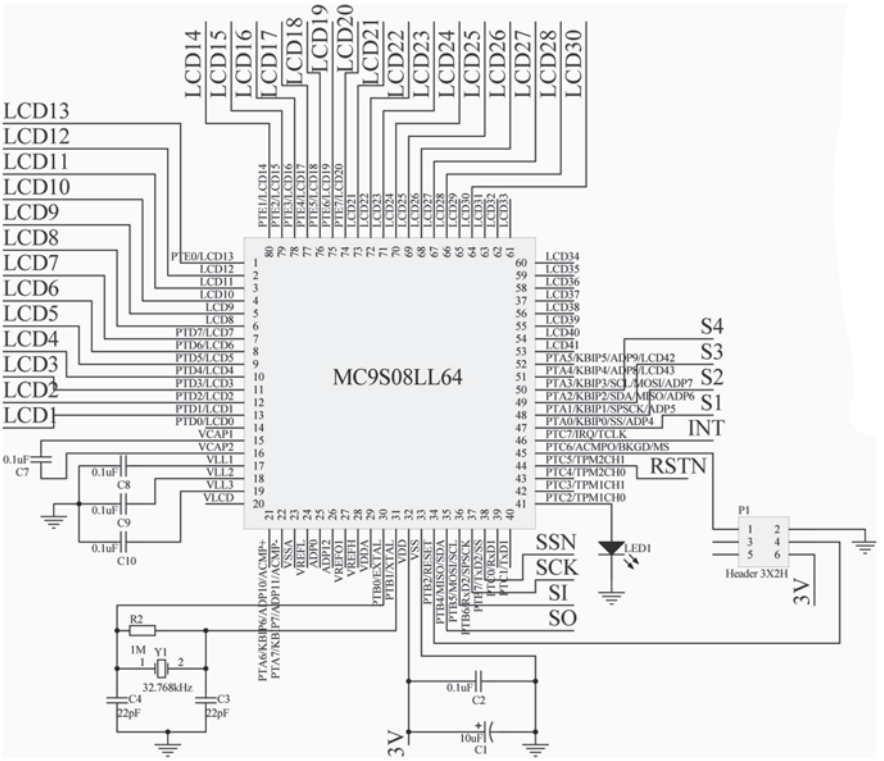


Fig. 83.4 Peripheral circuit of MCU. MCU micro control unit

by measuring the relative pulse width of the first wave. Then, TDC-GP22 can work out the time of flight (TOF) of upstream. Next, the same procedures will be executed again from the downstream ultrasonic transducer in order to measure the TOF of downstream. At the same time, TDC-GP22 also measures the real-time value of temperature through the platinum resistance temperature sensor; the temperature will be used as the basis for selecting the ultrasonic speed when calculating. With all the data being sent back to MCU, the result will be calculated and finally displayed on LCD.

As shown in Fig. 83.6, TOF measurement instruction will be sent twice per second and the temperature of fluid will be measured every 30 s. Importantly, the capability of first-wave detection gives MCU the possibility to judge whether bubbles exist in the propagation path. If any, the alarm lamp will be lighted. Until the strength of the received signals turns good, MCU starts calculations. In order to obtain better precision, the final instantaneous flow is acquired by averaging every five measurements.

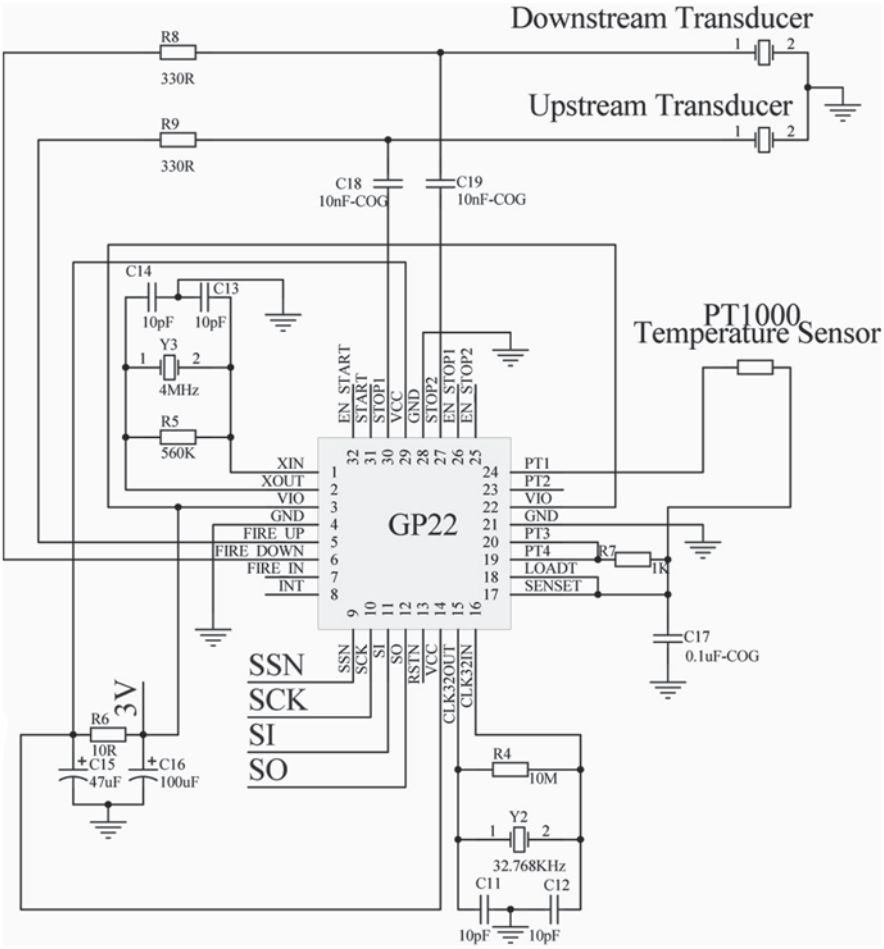


Fig. 83.5 Peripheral circuit of TDC-GP22 4.2 Software Design

83.5 Experimental Data and Calibration

In view of the difficulty of obtaining the high precision of instantaneous flow, a mass of identification experiments of instantaneous flow have been carried out after the design of measuring circuit. As a result, the error of experimental data does not reach the national standard of secondary meter. For the purpose of improving precision, the measuring data need to be calibrated before it is displayed on LCD. The algorithm of least square method is used to calibrate the measuring data [7]. On comparison between the experimental error before and after calibration as in Table 83.1, it can be found that the error after calibration has been decreased greatly and has met the national standard of secondary meter (less than 2 %).

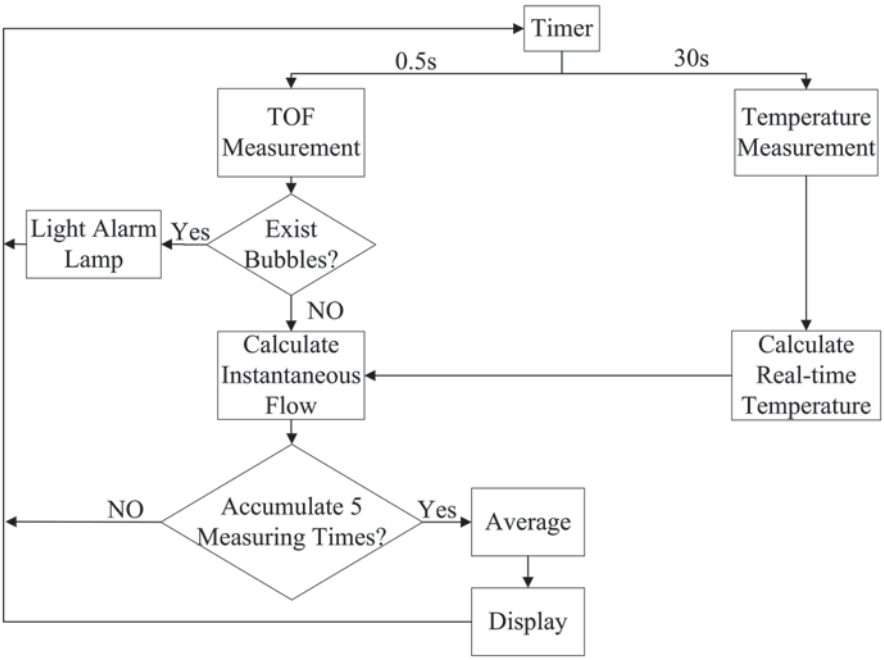


Fig. 83.6 Flow chart of main program

To demonstrate the superiority of this new V-reflection tubing, the experiments of traditional U-reflection circular tubing matched with the same measuring circuit have also been carried out and the experimental data are listed in Table 83.2. Comparing Tables 83.1 and 83.2, it can be easily found that the new V-reflection tubing can achieve better accuracy than the traditional circular U-reflection tubing.

Table 83.1 Experimental data of new V-reflection tubing before and after calibration

Standard value (m^3/h)	Original average value (m^3/h)	Original average error (%)	Calibrated value (m^3/h)	Calibrated error (%)
0.05	0.0434	13.20	0.050516257	1.033
0.1	0.0872	12.80	0.100184228	0.184
0.5	0.4182	16.36	0.503000283	0.600
1	0.8130	18.70	0.992543930	0.746
1.5	1.2238	18.41	1.505187993	0.346
2	1.6658	16.71	2.024980552	1.249
2.5	2.0348	18.61	2.483242429	0.670
3	2.446	18.47	2.996147304	0.128
3.5	2.8836	17.61	3.509912821	0.283
4	3.2798	18.01	3.997528279	0.062

Table 83.2 Experimental data of traditional U-reflection circular tubing

Standard value (m^3/h)	Final value (m^3/h)	Final error (%)
0.05	0.0486	2.76
0.1	0.1004	0.40
0.5	0.5052	1.04
1	0.9956	0.44
1.5	1.4738	1.76
2	1.9838	0.81
2.5	2.4618	1.53
3	2.9374	2.09
3.5	3.4287	2.04
4	3.9323	1.69

83.6 Conclusion

In conclusion, according to the method of time difference, a new measuring circuit based on GP22 and a new V-reflection measuring tubing are proposed for ultrasonic flow meter. The simulation result and the experimental data show that both of them have improved the precision of instantaneous flow measurement. As for future work, the parts of measurement method of temperature and temperature compensation can be further investigated on the basis of this proposed measuring circuit so as to develop a new type of ultrasonic heat meter.

References

1. Wang YH. Development of multi-path ultrasonic flow meter based on embedded system [C]. 2010 8th IEEE International Conference on Control and Automation. NJ: IEEE Computer Society; 2010. pp. 689–92.
2. Yang B, Cao L, Luo YP. High-speed and precise measurement for ultrasonic liquid flow metering based on a single FPGA [C]. 2009 IEEE Instrumentation and Measurement Technology Conference. NJ: IEEE Computer Society; 2009. pp. 309–12.
3. Shan HW. Study of micro power ultrasonic wave flow meter [C]. 2013 4th International Conference on Digital Manufacturing and Automation. Washington: IEEE Computer Society; 2013. pp. 1245–8.
4. Yu Y, Zong GH. Ultrasonic liquid flow meter for small pipes [J]. Rev Sci Instrum. 2012;83(2):026107.
5. Liu Q, Wang RD, Chen CG. Design and implementation of low-power time difference ultrasonic flow-meter [C]. 2011 International Conference on Mechanical Engineering and Technology. Heidelberg: Springer; 2012. pp. 469–75.
6. Yu Y, Zong GH. Design and simulation of an ultrasonic flow meter for thin pipe [C]. 2011 IEEE International Conference on Mechatronics and Automation. NJ: IEEE Computer Society; 2011. pp. 1115–9.
7. Li MW, Lv GS, Hu YG. Research on improving the accuracy of the ultrasonic flow-meter with time difference method [C]. International Conference on Electrical and Control Engineering. NJ: IEEE Computer Society; 2010. pp. 1704–7.

Chapter 84
Diagnosis Expert System Design of a Certain Ordnance Launch Circuit

Xiaoyan Qu, Baogang Li, Chengzhu Yao and Wei Pang

Abstract In order to solve the longtime problem of fault location and low efficiency of fault clearing for a certain ordnance launch circuit, a fault diagnosis expert system based on hybrid case-based reasoning (CBR) and rule-based reasoning (RBR) is designed. According to the structure of systematic fault tree and historical fault cases, fault repository is established. The inference mechanism based on hybrid CBR and RBR is adopted to the fault diagnosis. Practical application shows that the expert system can improve the efficiency of ordnance launch circuit fault diagnosis and reduce the fault-clearing strength of maintenance crew.

Keywords Fault diagnosis · Expert system · CBR · RBR

84.1 Introduction

Ordnance launch circuit is used to manage and control various weapons and ordnance carried by aircraft, which links air-to-air firepower control system, air-to-ground firepower control system, radar warning system, etc. with complicated signal relationship.

Rule-based reasoning (RBR) is the method through which domain experts and knowledge engineers communicate with each other to establish rule base [1–3]. It has advantages of expressing knowledge effectively and operating easily, but it is difficult to avoid the conflict and contradiction between the rules, which may be

X. Qu (✉) · B. Li · C. Yao · W. Pang
Naval Aeronautical and Astronautical University, 264001 Yantai, China
e-mail: 19804473@qq.com

B. Li
e-mail: 2844197381@qq.com

C. Yao
e-mail: 370775905@qq.com

W. Pang
e-mail: 719970894@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_84

caused by the problem of combinatorial explosion and mutual conflict in the process of inference. Therefore, case-based reasoning (CBR) is introduced to establish case base [4, 5]. CBR can improve case base through self-learning gradually in the reasoning process, which can overcome the shortcomings of RBR. But CBR cannot express the knowledge which can be understood accurately by human, and CBR has limited successful case in the case base.

Based on the above analysis, this chapter designs a fault diagnosis expert system based on hybrid CBR and RBR for a certain ordnance launch circuit, which not only can make up for the disadvantages of rule extraction difficult and few case quantity but also can keep the advantages of high efficiency in the case of reasoning and easy to operate. The proposed method improves the efficiency and accuracy of the fault diagnosis for ordnance launch circuit.

84.2 The Structure of Fault Diagnosis Expert System

For a certain ordnance launch circuit, the system uses hybrid CBR and RBR to design fault diagnosis expert system as shown in Fig. 84.1, which includes case base module, rule base module, dynamic database module, reasoning module, knowledge base management module, explanation module, knowledge acquisition module, and man–machine interface.

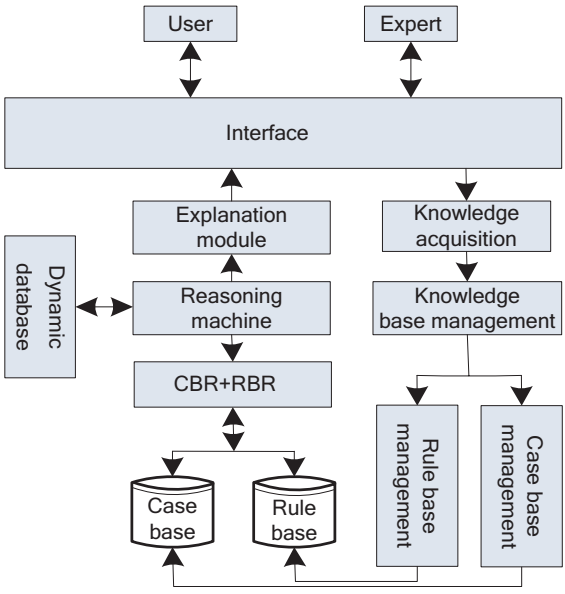


Fig. 84.1 Hybrid fault diagnosis expert system. *CBR* case-based reasoning, *RBR* rule-based reasoning

84.3 Case Base Module

In fault diagnosis of ordnance launch circuit, one case is a complete solution of one failure phenomenon. This solution can be used for the same or similar fault diagnosis. According to characteristics of this system, using case description method of basic properties for ordnance launch circuit fault [6], namely case representation, is composed of six parts: case number, case name, case creation time, case symptoms, failure reason, and troubleshooting. According to the six parts, make corresponding data table to store case information. Take ordnance launch circuit bus signal conversion board fault as an example, to establish case, as shown in Table 84.1.

84.3.1 Rule Base Module

By analyzing structure and principle of ordnance launch circuit, the fault tree is built. The rule knowledge is represented by production-based model and is stored in database in the form of table. Its general form is:

If <premise>, then <conclusion> <credibility>.

The conversion principle for fault tree converting into rule knowledge base is:

1. In fault tree, the logical "AND gate" is equivalent to rule 1.
2. In fault tree, the logical "OR gate" is equivalent to multiple rules. Take bus signal conversion board damage, for example.

Table 84.1 Bus signal conversion board fault case

Case no	00010	
Case name	Bus signal conversion board damage causes no output of power module	
Create time	D2012-01-20T10:24	
Case symptoms	Side 1: measuring voltage of aircraft to missile power signal no output	Weight: 0.9
	Side 2: measuring voltage of aircraft to missile power signal too high	Weight: 0.6
	Side 3: measuring voltage of aircraft to missile power signal too small	Weight: 0.7
	Side 4: measuring voltage of aircraft to missile power signal over 30 s	Weight: 0.9
	Side 5: three-phase AC output is normal	Weight: 0.9
Fault reason	Bus signal conversion board damage	
Repair method	Measure bus signal conversion board output; to check the fault unit; replace bus signal conversion board	

AC alternating current

Table 84.2 Structure representation of rule knowledge base

Item	Premise	Conclusion	Credibility
1	X_1, X_2, X_3	Bus signal conversion board fault	0.90
2	X_1, X_2, X_4	Bus signal conversion board fault	0.85

To convert the fault tree of bus signal conversion board damage into rule knowledge base, $X_i (i = 1,2,3,...)$ is assumed to represent basic event. The structure is shown in Table 84.2.

84.3.2 Dynamic Database Module

Dynamic database mainly stores some data record during system operation and original data character in the field of diagnosis, in the form table storage. It mainly includes nominal value table, measured value table, fault phenomenon table, and diagnosis result table. Nominal value table mainly stores various parameters value under launch circuit normal status, measured value table stores actual value of measured parameters, fault phenomenon table stores project with fault in launch circuit, and diagnosis result table stores diagnosis result and corresponding repair advice. In which, measured table is provided by test program or manual work, then compared with nominal value in the initial stage of expert system operation, and then the fault phenomenon table is obtained. The fault phenomenon table proceeds hybrid reasoning through case base and rule base, gets diagnosis result to put in diagnostic result table.

84.3.3 Reasoning Module

1. Hybrid Reasoning Mechanism

The system adopts the reasoning mechanism of hybrid CBR and RBR. For a new problem, start CBR first, if the similar problem solving cannot be found previously, and then turn to RBR. Inference control process is shown in Fig. 84.2.

2. CBR Mechanism

CBR is generally divided into the following four steps:

1. Input requirements, parameters, and other relevant information of the problem to be solved.
2. According to the requirements and initial conditions, retrieve similar cases from the case base.

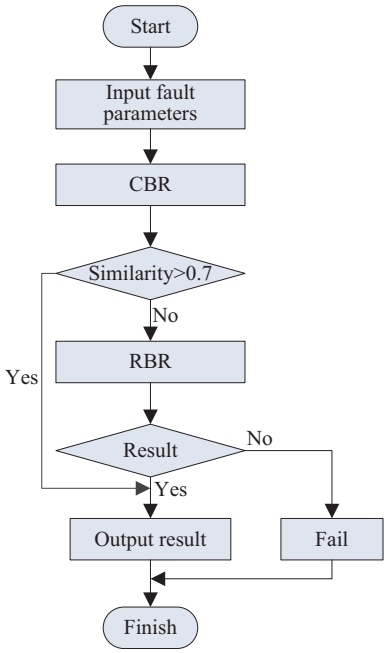


Fig. 84.2 System inference controlling flow. *RBR* ruled-based reasoning, *CBR* case-based reasoning

3. To find the most similar case from the similar case, or by a number of case mix, to form the target problem solution, to meet current requirements through object scheme modification.
4. When problem is solved, current solution can be stored into case base as a new case; reasoning process is shown in Fig. 84.3.

The most important for CBR mechanism is the establishment of retrieval mechanism, and finding out the most similar case with current task case. Common retrieval algorithms are nearest strategy method, inductive indexing method, and

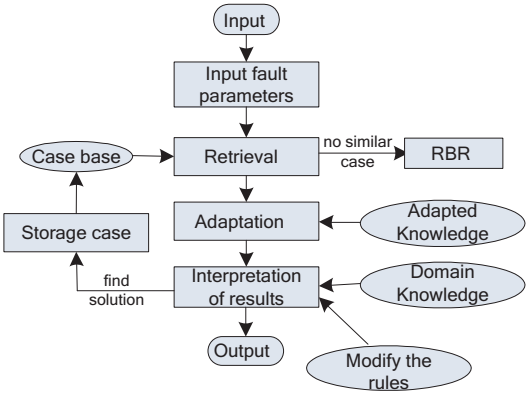


Fig. 84.3 Reasoning process based on CBR. *RBR* rule-based reasoning, *CBR* case-based reasoning

knowledge guidance method [7]. As ordnance launch circuit test has strict timing relationship, this chapter adopts the nearest strategy approach, and algorithm is as follows.

Assume $F(F_1, F_2, F_3, \ldots, F_m)$ is the results of ordnance launch circuit measure parameters currently input by user, m is the total number of ordnance launch circuit measure parameters. If the ith parameter is qualified, then $F_i = 0$, otherwise, $F_i = 1$, and $i = 1, 2, \ldots, m$.

Assume $P(P_1, P_2, P_3, \ldots, P_m)$ is a case in case base. Method to get P_i is same as F_i. $W(W_1, W_2, W_3, \ldots, W_m)$ is the weight vector of parameter. According to ordnance launch circuit test parameter setting, use bigger weight value for parameter which plays an important role in fault location; the weight value can be adjusted according to the actual using situation. Therefore, after determining weight coefficient W_i value of each parameter, you can determine current input problem and matching degree value $S(F, P)$ of the case in case base.

$$S(F,P) = \frac{\sum_{i=1}^{m} W_i \times S(F_i, P_i)}{max\ (N_P, N_f)} \tag{1}$$

In the formula, $S(F_i, P_i)$ uses "AND" operation of F_i and P_i, $N_P = \sum_{i=1}^{m} P_i$ and $N_f = \sum_{i=1}^{m} F_i$ represent the number of parameter group and fault parameter of case database, respectively.

3. RBR Mechanism

According to the sign of running state of unit under test, RBR mechanism uses knowledge of diagnosis database to get reasoning strategy by comparison, reasoning, and diagnosis for historical database.

Reasoning strategy is divided into forward reasoning, backward reasoning, and hybrid reasoning. As ordnance launch circuit fault parameters are given, namely start from given facts, using the knowledge in knowledge base, it is to find all fault reasons which can be deduced; therefore, it uses forward inference.

Forward reasoning based on rules uses fact-driven mode for reasoning under the premise of rule knowledge existence. The system matches fault symptom information to be diagnosed and prerequisite rules of rule base, and if the matching is successful, then use the conclusion of knowledge block as intermediate results, and use the intermediate results to continue to match rules in knowledge base, until you get ordnance launch circuit fault judgment results.

In the process of RBR, if only one rule matching is successful, it is the most ideal situation, but if there are two or more rules that match successfully, they become campaign rules. At this time, reasoning mechanism resolves conflict through using depth-first searching algorithm. The algorithm is that from the starting node along a route to the left down until finding a termination node or

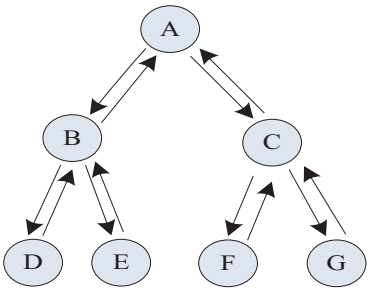


Fig. 84.4 Depth-first search algorithm

target. If it gets to termination node, and the node is not a goal, then it returns to the node of last layer, goes down along the node of next layer's left until the target or termination is approached. If all the nodes of the node's next layer are already searched, then it returns to the last layer. Duplicate this process until you find the target or retrieve the last node in the searching space. Depth-first searching strategy is shown in Fig. 84.4. If traversal of nodes is in F, traverse this picture by order ABDBEBACF.

Ordnance launch circuit fault diagnosis expert system uses forward reasoning and depth-first reasoning algorithm to conflict resolution to obtain answer [8]. First, input fault phenomenon and then conduct rule matching; if two or more rules match, then take depth priority reasoning algorithm to select one to perform according to the high credibility principle. Inference flowchart based on rules is shown in Fig. 84.5.

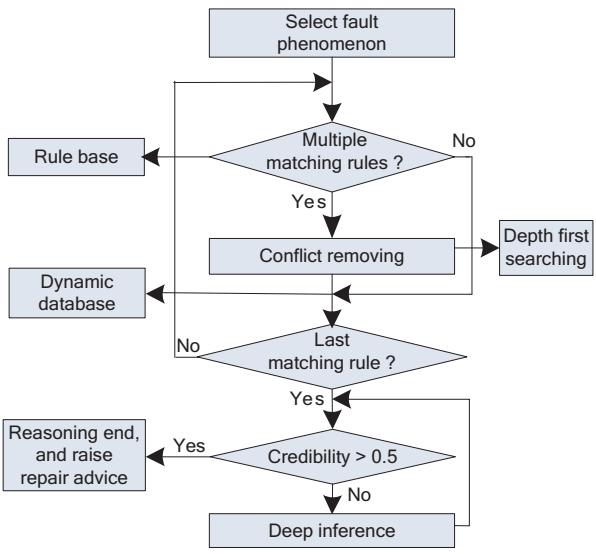


Fig. 84.5 Inference flowchart based on RBR. *RBR* rule-based reasoning

84.3.4 Other Module

1. Knowledge base management module: creation, delete, query, and update of knowledge base.
2. Explanation module: answers questions raised by the user using preset text method, and gives reasonable explanation for solving process and the results of the expert system, and forms diagnostic report, gives the final diagnosis result and repair advice.
3. Man–machine interface: interacts between user and expert system, which includes two parts: input and output. One part identifies and understands informationand then inputs to system; the other part transfers the results produced by expert system into form which human can accept.

84.4 Achievement

This chapter takes Visual.net as the development platform, Access 2003 as the database structure, and ActiveX Data Objects (ADO) as the database access technology to develop a hybrid fault diagnosis expert system for a certain ordnance launch circuit.

1. Login module: it is divided into two login modes of user and expert, and provides to users and experts to use.
2. Expert module: expert maintains rule base based on fault tree model.
3. User module: user get preliminary diagnosis results through inputting fault phenomenon, and at the same time combining with other fault phenomenon and the results detection equipment to conduct deep fault diagnosis. The diagnostic results can be saved, searched, and printed.

84.5 Conclusion

This chapter designs a fault diagnosis expert system for a certain ordnance launch circuit based on hybrid CBR and RBR. The proposed method overcomes difficulties such as difficult ruled knowledge acquisition and low reasoning efficiency in single RBR inference. This design improves diagnosis efficiency and diagnosis result reliability of system, and reduces the fault-clearing strength of maintenance crew.

References

1. Rissland L, Kalak BS. CABARET: rule interpretation in hybrid architecture. Int J Man-Mac Stud. 1991;34(6):839–87.
2. Branting LK. Building explanations from rules and structured cases. Int J Man-Mac Stud. 1991;34(6):797–837.
3. Crone N, An A, Chan C. Rule-induction and case-based reasoning: hybrid architecture appear advantageous. IEEE Trans Knowl Data Eng. 1999;11(1):166–74
4. Kolodner JL. An introduction to case based reasoning. Artif Intell Rev. 1992;6(1):3–34
5. Modt A. Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun. 1994;7(1):39–59.
6. Li Q, Fang J. Knowledge representation of planning based on RBR and CBR. Comput Eng Des. 2009;30(22):5166–74.
7. Babaka O, Whar SY. Case-based reasoning and decision support systems. IEEE International Conference on Intelligent Processing Systems, IEEE. 1997: 1532–1536.
8. Ma D, Ma L, Lv X. A fault diagnosis expert system for a certain airplane's missile-launch channel. Comput Meas Control. 2011;19(11):2705–7.

Chapter 85
Arcless Interruption Method of an Alternating Current Contactor

Ming Zong, Tianyu Dong and Zhengdi Cao

Abstract The arc available during opening and closing of alternating current (AC) contactors' contacts is the main factor influencing a contactor's life. In this chapter, the modular asynchronous contactors which can control each contact individually are proposed. When one of the three-phase AC currents reaches zero point, the contactor breaks the contacts. This can achieve the contactor break off without arc. It is a new idea in the field of structural design and control of contactors. This chapter analyzes the method that contactor achieves breaking without the arc; and according to the characteristics of the three-phase AC circuits, when the first phase of the three-phase AC contacts off, the detailed discussion and analysis are made to probe into the relationship of currents change between the other two phases of the three-phase AC circuits, deduce the formula about the time from the first phase of the three-phase AC contacts off to the other phases of the three-phase AC contacts off. And the accuracy of this formula is verified in matrix laboratory (MATLAB).

Keywords Contactor · Arcless interruption · Asynchronous modular · Simulink modeling

85.1 Introduction

The arc available during the opening and the closing of alternating current (AC) contactors' contacts is the main factor which may influence a contactor's life. This chapter presents the asynchronous modular contactors for studying new ways so that the contactor can open without arc. The asynchronous modular means a

M. Zong (✉) · T. Dong · Z. Cao
School of Electrical Engineering, Shenyang University of Technology, 110870 Shenyang, China
e-mail: ming_zong@163.com

T. Dong
e-mail: tianyu_dong@126.com

Z. Cao
e-mail: caozdd@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_85

combination of three-pole contactor contained in the three independent contacts system. Combined with previous asynchronous, the difference point is the structure of the contact. Previous asynchronous modular contactor is that three contactors control three-phase power source, respectively. The new structure is that three separate contacts are applied to a contactor. Because the contactor adopts a new split-phase structure, it can decrease the spring force and control contact easier so that the volume will be smaller. Zhihong Xu [1] improved the structure of the AC contactor. It changed the structure of contact system. The three-phase contacts have a different distance. The first phase opening range is much longer than the others'. By controlling the breaking time of the first phase, the contactor can achieve three-phase contacts system with the control of zero current interruption [1]. Capp proposed the theory of arc electric potential model. This theory pointed out that if the opening range of contacts can be kept by 0.1 mm, two electrodes would receive equal energy from the arc. Thus, the erosion of contacts is the minimum [2]. Kiely designed an electromagnetic system with a power electronic circuit. Contactor received a small excitation current in the first half wave of power supply. Let the speed of moving parts reach a nearly constant 10 mm/s. The contact distance can reach 0.1 mm at the end of the first half wave. After the first half wave, the contactor received a large excitation current and the contact broke faster. At present, the application of this technology is limited to small capacity of AC contactor [3]. Now there are few researches in terms of the split-phase and the arcless interruption.

85.2 Working Principle of the Asynchronous Modular Contactor

The electric power production, transmission, and power supply way in the power system mostly use the three-phase system at present. The three-phase power supply, three-phase system load, and three-phase transmission lines constitute the three-phase circuit [4]. Asynchronous modular AC contactor proposed in this chapter controls the contactor breaking process through the split phase. When the contactor needs to open the contact, the first step is to open first phase contact in the next passing zero point, and then open the other two phase contact after the predefined time; thus, the three contacts may achieve the goal of arcless interruption. However, due to different working environments of the contactor, the predefined time is uncertain. The time is calculated by examining certain conditions.

This is the analysis of the three-phase symmetrical load circuit with Y–Y connection. The three-phase symmetrical circuit points to the three-phase power supply that is symmetrical, the three back line impedance that is equal, and the symmetrical three-phase load that is equal. As to the Y-shaped three-phase power supply, the phase voltage are $\dot{U}_A$, $\dot{U}_B$, $\dot{U}_C$. The corresponding line voltage, respectively, are $\dot{U}_{AB}$, $\dot{U}_{BC}$, $\dot{U}_{CA}$, and the phase current are $\dot{I}_A$, $\dot{I}_B$, $\dot{I}_C$. The phase current is equal to the line current.

$$\begin{aligned}\dot{U}_A &= U_m\angle 0^\circ \\ \dot{U}_B &= U_m\angle -120^\circ \\ \dot{U}_C &= U_m\angle -240^\circ = U_m\angle 120^\circ.\end{aligned} \tag{1}$$

Then, according to the Kirchhoff's voltage law (KVL), the line voltage is:

$$\dot{U}_{BC} = \dot{U}_B - \dot{U}_C = (1-a^2)\dot{U}_B = \sqrt{3}\dot{U}_B\angle 30^\circ = \sqrt{3}U_m\angle -90^\circ. \tag{2}$$

Assuming the symmetrical three-phase load is $Z = z\angle\phi$, the phase current is as follows:

$$\begin{aligned}\dot{I}_A &= \frac{\dot{U}_A}{Z} = \frac{U_m}{z}\angle -\phi \\ \dot{I}_B &= \frac{\dot{U}_B}{Z} = \frac{U_m}{z}\angle -120^\circ - \phi \\ \dot{I}_C &= \frac{\dot{U}_C}{Z} = \frac{U_m}{z}\angle 120^\circ - \phi.\end{aligned} \tag{3}$$

If the A-phase looks as the first opening phase, after the A-phase contact break at the zero point, the A-phase line current is $\dot{I}_A = I_A\sin(\omega t - \phi_A) = 0$, that is, $\omega t_0 = \phi_A$, and the loads of B-phase and C-phase are series. At this time, set the line current as:

$$\dot{I}_B = \frac{\dot{U}_{BC}}{2Z} = \frac{\sqrt{3}U_m}{2z}\angle -90^\circ - \phi. \tag{4}$$

Similarly, $\dot{I}_C$ and $\dot{I}_B$ are the opposite. Thus, we come to a conclusion: In case of outage of the A-phase power, the amplitude of the other phase contacts' line current is equal, the original amplitude of $\sqrt{3}/2$ times. The phase contrast between the B-phase and the C-phase is 180°. At the moment when $\omega t = \phi$, A-phase breaking as the starting moment, there are B and C line current expression of the two phases shown as follows:

$$\begin{aligned}\dot{I}_B &= \frac{\sqrt{3}}{2}I_m\angle -90^\circ - \phi \\ \dot{I}_C &= \frac{\sqrt{3}}{2}I_m\angle 90^\circ - \phi.\end{aligned} \tag{5}$$

Based on the abovesaid analysis, the result is as to the symmetrical three-phase circuit with Y–Y connection, and in case of outage of the A-phase power, the amplitude of the other phase contacts' line current is equal. The phase contrast between

the B-phase and the C-phase is 180°. And upon the outage of A-phase power, the current of B-phase and C-phase become zero for the first time after a quarter of the period. The three-phase power supply frequency is 50 Hz, so period of 20 ms, and thus a quarter of a period is 5 ms.

According to the actual work situation of AC contactor, the contacts need to preset at 5 ms off. Upon breaking for several times, according to the arc current sampling signal provided by the sensor, the time is adjusted. It can make the moment of the contact breaking to return to the zero point of current and realize the adaptive control. As the best breaking area is small, the step changed every time is 0.2 ms. The way is the two-way movement. If there is the arc at the same direction upon changes for three times, restore the initial value and move to another direction. As circulation, it can guarantee the contactor breaking in zero point of current [5]. This way can make the contacts to break off quickly and accurately, and reduce the arc of breaking contactor, which can be realized with arcless interruption.

85.3 Verification of Simulation

As to the three-phase circuit with Y–Y connection, in case of outage of the A-phase power, the amplitude of the other phase contacts' line current is equal. The phase contrast between the B-phase and the C-phase is 180°. In other words, upon the outage of the A-phase power, the current of B-phase and C-phase become zero at the first time in the ***t*** ms. The symmetrical three-phase circuit model is established by Simulink as shown in Fig. 85.1. The three-phase voltage source is, respectively, A-phase

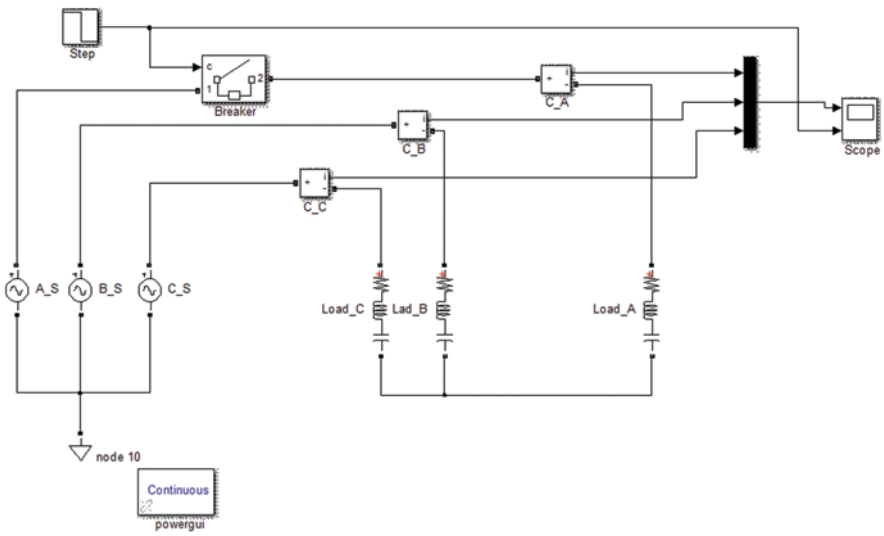


Fig. 85.1 Simulink model of the symmetrical three-phase circuit

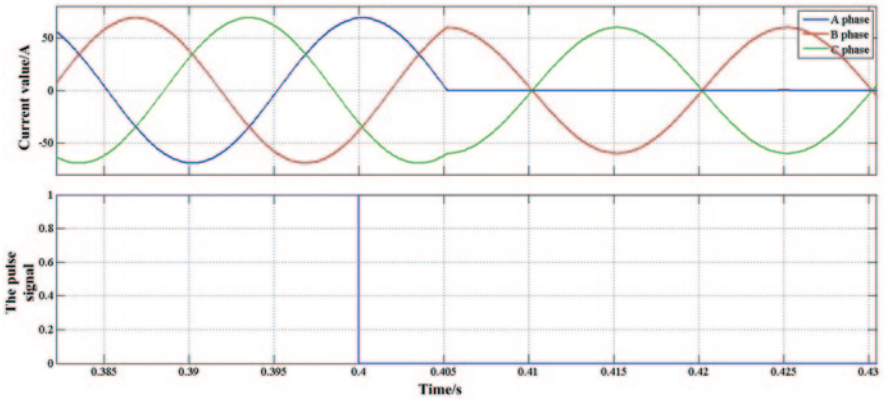


Fig. 85.2 Result of Simulink software simulation

is $220\sin(\omega t)$, B-phase is $220\sin(\omega t - 120^{\circ})$, and C-phase is $220\sin(\omega t + 120^{\circ})$. The power frequency is 50 Hz. In order to facilitate the calculation, the three-phase load set is that the three-phase symmetrical load in this chapter. The breaker module can be equal to the breaker. When the input of control port is the "0" signal, the breaker breaks at the zero point of current.

When the input of control port is the "1" signal, the breaker closes the circuit. Its external input is step signal. When the time is 0.4 s, the signal is shut off. This can make the A-phase circuit open in the current zero of nearby 0.4 s. When the three-phase electric power is supplied, the simulation results are obtained as follows (Fig. 85.2):

Through the formula of the last section, it can be obtained: Upon the outage of the A-phase power outage, the current of B-phase and C-phase become zero at the first time at 5 ms. From the simulation results, the time is from 0.405 to 0.41 s, this time is 0.005 s, which is consistent with the result of the formula deduced.

85.4 Conclusion

This chapter presents an interruption method of asynchronous modular contactors. As this new contactor structure is different from the conventional contactor, the breaking methods are also different from the conventional contactor. Due to the uncertain working conditions, the contactor breaking conditions are also different. By means of the formula derivation and conditions analysis, obtain the contactor's breaking methods and breaking time, and the simulation can prove that it is correct. As this chapter is only for simulation instead of verification by the test, it should be optimized in the future work.

Acknowledgments This work is supported by the National Natural Science Foundation of China (Grant No.51177105).

References

1. Xu Z. Zero current interrupt control of intelligent AC contactor. Adv Technol Electr Eng En. 2002;21(4):55 (In Chinese).
2. Capp B. The power balance in electrode dominated arcs with a tungsten anode and a cadmium or zinc cathode in nitrogen. J Phys D: Appl Phys. 1972;5(12):46.
3. Kiely JH, Nouri H, Kalvelage F, et al. Development of an application specific integrated circuit for reduction of contact bounce in three-phase contactors//46th IEEE Holm conference on electrical contact. Chicago:IEEE, 2000:120–9.
4. Lin X. Simulation of three-phase circuit based on multisim [J]. Mod Electron Tech. 2002;(4):131 [In Chinese].
5. Zheng X. Analysis and implementation of zero-current breaking adaptive control in intelligent AC contactor [J]. Adv Technol Electr Eng En. 2005;1.24(3):77 [In Chinese].

Chapter 86
Analysis of 750-kV High-Resistance Operation

Jingjing Zheng, Yong Yang and Guojin Yang

Abstract The controllable high resistance of the 750-kV transformer substation bus in D is the first controlled high resistance of the 750-kV voltage level, which is due to the lack of actual operating experience. The controllable high resistance in airdrop and capacity adjustment will produce harmonic voltage or current at different levels and of different frequencies, which may threaten the stability and safety of system operation and equipment seriously. In this chapter, the controllable high resistance during airdrop and run adjustment are studied by recording the electrical quantities and analyzing the harmonic. The effect of the second harmonic to the high-voltage direct current (HVDC) system and the hazards of the direct current (DC) component to the reactor are adjusted to obtain the harmonic characteristics of the controllable high resistance of the 750-kV transformer substation bus in D. Based on the above analysis, some recommendations are made to maintain the controllable high resistance of the 750-kV transformer substation bus in D so as to alleviate the frequent switch of reactive power compensation equipment as a result of high voltage and stochastic volatility.

Keywords Controllable high resistance · Airdrop · Inrush current · Second harmonic · DC component

86.1 Introduction

The bus controlled high voltage reactor of 750kV Dunhuang substation is Chinese first controlled high voltage reactor of 750kV level, which is on behalf of the international leading level in this field. This equipment was developed combining with the domestic and foreign practice of controlled high voltage reactor and the situation of 750kV power grid. It is set beside the 750-kV bus of D substation with

J. Zheng (✉) · Y. Yang
Gansu Electric Power Research Institute, 730050 Lanzhou, China
e-mail: yy8801@126.com

G. Yang
Economic Research Institute of Qinghai Province, 810008 Xining, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_86

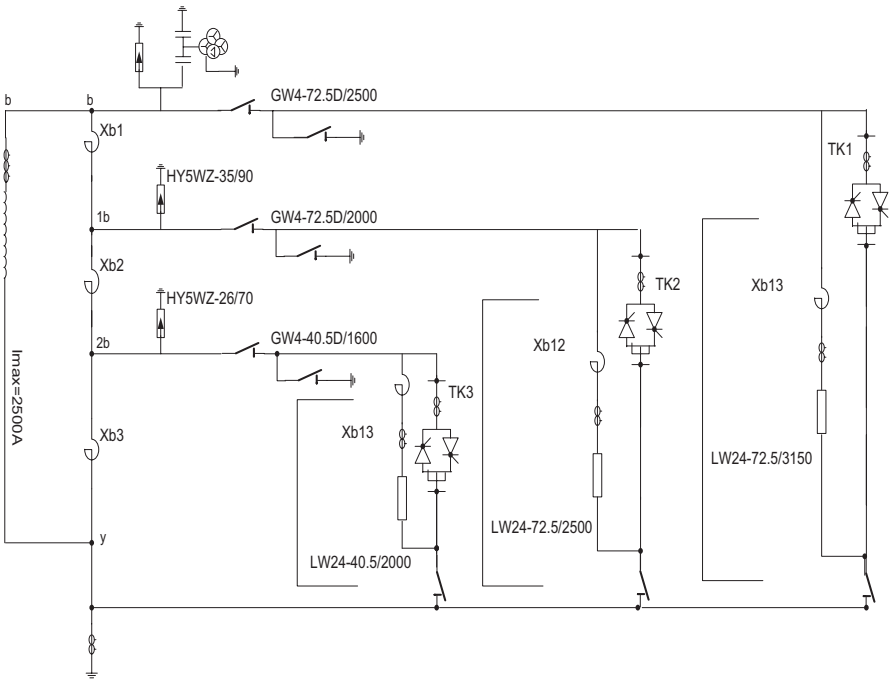


Fig. 86.1 A 750-kV controllable high-resistant secondary structure

the capacity of 300 and is divided into four controllable levels of 100, 75, 50, and 25 %. The operation of this system will be conducive to resolve the conflict between the Hexi 750-kV grid reactive adjustment and the overvoltage limit, and alleviate the problems such as long-distance transmission of the Xibei grid, intermittence of wind energy resources, and voltage stability problems resulting from the volatility; besides, it may reduce the transmission loss, improve the transmission capacity and grid stable safety, and provide advantageous conditions for Hexi wind power outgoing. A 750-kV controllable high-resistant secondary structure is shown in Fig. 86.1.

86.2 Inrush Current Analysis of Controllable High Resistance

With the rapid development of grid technology, long distance, and huge capacity, the alternating current (AC)/direct current (DC) hybrid power operation is bound to be a trend. At present, the Xibei grid DC transmission line is booming greatly. The AC and DC control and protection coordination has an important influence on the safety and stability of grid. The DC frequency stabilization and the secondary harmonic protection are the basic protection functions of the AC and DC

hybrid system. Actually, these two kinds of protection can be generally called the DC harmonic protection or the fundamental frequency protection. When dropping, the large power transformer will produce different sizes of excitation inrush current, in which the second harmonic occupies a large proportion.

In the actual operation, when the main transformer of 750 kV is subject to airdrop, it will cause an accident when the second harmonic causing the DC converter station bipolar is blocked. In the case of load loss, a serious threat will also be posed to the safety of power grid. The D-controllable high-resistant Y/delta junction method has almost the basically consistent electromagnetic structure and transformer. There must be a certain inrush current in the case of airdrop. As the DC power transmission project in the northwest grid grows rapidly, when the inrush current is caused by the controllable high resistance of 750 kV, it will not only affect their own protection but also be more likely to pose a threat to the normal operation of the DC transmission lines.

Factors affecting transformer inrush current shall include the following content:

The power voltage. Upon the closing of the transformer, the amplitude of the iron core forcing flux $\Phi_m = U_m/W_m$; therefore, the higher the voltage, the bigger the Φ_m, and the greater inrush current will be.

The closing angle α. When the closing angle $\alpha = 0$, $\Phi_m \cos \alpha$ reaches its minimum inrush current, and when $\alpha = 90$, the inrush current reaches its minimum.

The remanence B_s. Before closing, the greater the remanence in the transformer core is, the greater the inrush current [1] will be.

Upon testing the large transformer winding DC resistance, it can generate a large remanent in the core and cause a dramatic increase of inrush current of transformer when airdropping [1–4] with a huge impact on the grid; at the same time, the even harmonics which give priority to the harmonic will be a direct threat to the safe and stable operation of the AC/DC system [5–9].

The study of the second harmonic influence on the controllable high resistance when airdropping is of great significance to ensure the safe and stable operation of the northwest power grid DC transmission line, and improve the reliability and stability of the system.

Figures 86.2 and 86.3 show, respectively, the wave recording of three-phase voltage and three-phase current beside high voltage when the controllable high resistance of 750 kV is charged for the first time. From the wave record, it can be found that the inrush current stands most in the three-phase current; thus, it has affected the three-phase voltage waveform. Among them, it is the DC distribution leading the voltage, and the current waveform leans to one side of the time axis, and it is mainly the second harmonic distribution that causes the waveform distortion. Comparing the controllable high resistance and voltage and current harmonic analysis of the main transformer of 750 kV in the case of airdrop, the second harmonic changing trend can be found, which is shown in Fig. 86.4.

In Fig. 86.4, when the controllable high resistance of 750-kV main transformer is airdropped, the amplitude of second harmonic change or changing trends is very similar; therefore, it is suggested that the controllable high-resistant winding should

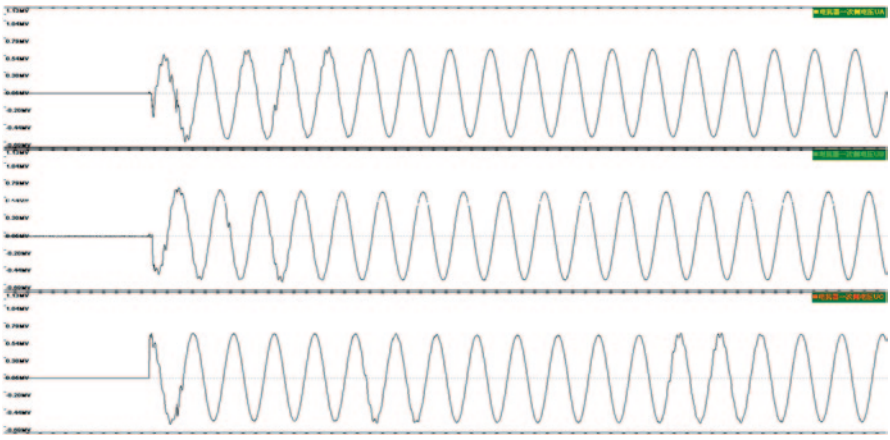


Fig. 86.2 High-voltage three-phase voltage waveform when the controllable high resistance of 750 kV is first charged

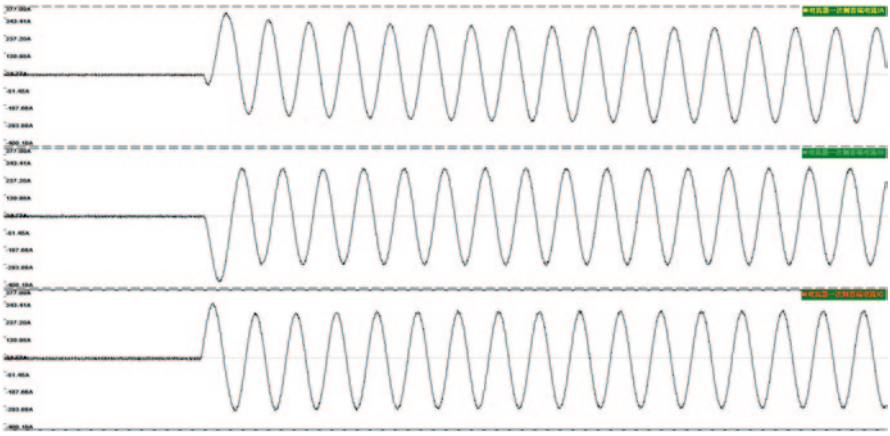


Fig. 86.3 Side high-voltage three-phase voltage waveform when controllable high resistance of 750 kV is first charged

take appropriate measures after the DC resistance test in order to reduce the inrush current and the second harmonic component.

Besides, in the analysis of the controllable high resistance of 750 kV and three main transformers in the case of airdrop, it has been found that the second harmonic inrush current component decays relatively quickly; usually, it reduces to a relatively lower level in a frequency attenuation (ms) of 50 or so. Therefore, the harmonic (100 Hz) protection of the DC transmission system is set with full consideration taken meanwhile so as to avoid the large transformer and reactor decay time of second harmonic of nearby nodes.

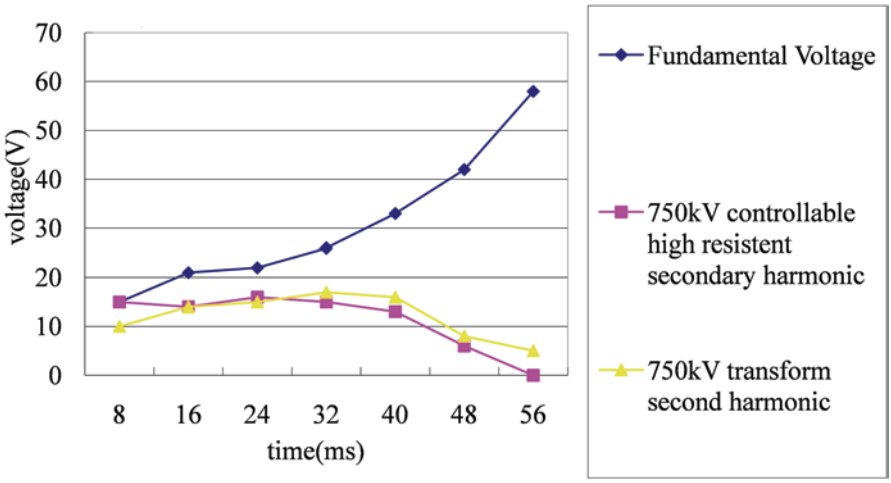


Fig. 86.4 Secondary harmonic changing trend when the controllable high resistance and main transformer airdrops

86.3 Harmonic Analysis of Controllable High-Resistant Adjustment

Bypass circuit breakers are paralleled with the classification of adjusting thyristor valves at the secondary side of a 750-kV controllable high-voltage reactor, thus ensuring the controllable high-voltage reactor still works when there is some failure in the valve control system. At the same time, through the following means, such as the cooperation of thyristor valves and switches, cascading reactors in switch branches, the electrical condition is improved [10–11].

As the 750-kV substation in D in the Jiuquan wind power base is located at the central position, and the outgoing of wind load changes frequently; therefore, the regulation of 750-kV controllable high resistance will be more frequent. In the adjustment of controllable high resistance, the thyristor valve is seated or in full conduction state; therefore, there is no harmonic when the regulating process is completed. But in the process of adjustment, there are still relatively obvious harmonic components. Tables 86.1 and 86.2 list the maximum components of the harmonic and DC distribution during the adjustment.

According to the statistical analysis of Tables 86.1 and 86.2, the following rules can be found.

The side harmonic of controllable high resistance and low voltage is greater; relatively, it can be neglected for the high-side harmonic voltage. When the controllable high-resistant capacity is adjusted from 75 to 75%, or from 100 to 75%, the harmonic current on both sides stands most.

In addition, from the analysis on the current of the high-voltage side and low-voltage side, it is found that harmonic current attenuates very quickly two times,

Table 86.1 Controllable high resistance in lifting capacity adjustment in the process of harmonic maximum content in statistics

		25→50%	50→75%	75→100%
High-side voltage	DC component maximum	0.28	0.28	0.13
	Secondary harmonic maximum	0.26	0.25	0.29
	Third harmonic maximum	0.42	0.13	0.33
Low-side voltage	DC component maximum	15.37	15.71	33.00
	Secondary harmonic maximum	10.10	32.62	55.21
	Third harmonic maximum	6.52	17.28	39.03
DC-side voltage	DC component maximum	69.13	24.64	17.88
	Secondary harmonic maximum	28.06	14.03	10.12
	Third harmonic maximum	16.54	8.94	5.71
Low current side	DC component maximum	62.41	20.13	20.88
	Secondary harmonic maximum	28.23	14.14	9.84
	Third harmonic maximum	16.85	8.90	5.77

DC direct current

Table 86.2 Statistics of maximum harmonic content in the adjustment of controllable high resistance

		25←50%	50←75%	75←100%
High-side voltage	DC component maximum	0.30	0.24	0.32
	Secondary harmonic maximum	0.43	0.25	0.49
	Third harmonic maximum	0.36	0.34	0.41
Low-side voltage	DC component maximum	10.17	18.98	30.93
	Secondary harmonic maximum	17.50	29.53	56.03
	Third harmonic maximum	8.71	16.82	23.97
High-voltage-side current	DC component maximum	35.25	18.65	17.72
	Secondary harmonic maximum	21.80	13.96	10.37
	Third harmonic maximum	9.83	8.69	6.37
Low-voltage-side current	DC component maximum	27.85	22.97	14.59
	Secondary harmonic maximum	21.63	13.86	10.39
	Third harmonic maximum	14.08	8.53	6.29

DC direct current

three times, or higher, which can attenuate to a relatively low level in the time of one frequency or so. But the DC current component attenuates slower, usually, it takes the time of more than ten cycles, or even longer to attenuate lower levels.

When the direct current flows into the (controllable high resistance) transformer windings, there will be serious magnetic saturation to the transformer (controllable high resistance), which will result in a huge increase in inrush current and harmonic current, and may endanger the safe running of the equipment itself.

Therefore, it is recommended that the controllable high resistance should run in coordination with the main transformer of 750 kV at low-voltage-side capacity and reactance by making it more likely to run in the middle positon of 50 or 75%

capacity. In this way, certain adjustment capacity margin can be remained for controllable high resistance, meanwhile, the disadvantageous of large harmonic current of controllable high resistance can also be avoided.

86.4 Conclusion

The operation of D-controllable high resistance of 750 kV may alleviate the frequent switch of the reactive power compensation equipment to some extent which is caused by the high voltage and stochastic volatility. At present, there are few application examples of the controllable high-voltage shunt reactor. In the future operation and maintenance, with the accumulation of the experience in this regard, it will give full play to its fast reactive power compensation and voltage control ability of controllable high resistance.

References

1. Jiangsu electric power company. Electrical system relay protection principle and practical technology [M]. Beijing: China electric power press; 2006. pp. 15–20. (In Chinese).
2. Zhang B, Yin X. Power system relay protection [M]. Beijing: China electric power press; 2005. pp. 46–50. (In Chinese).
3. Wang W. The main electrical equipment relay protection principle and application [M]. Beijing: China electric power press; 1998. pp. 35–8. (In Chinese).
4. Li Z. Damage, regularity and restraint of higher harmonic power in network [J]. Electr Switch. 2009;2:72–6. (In Chinese).
5. Fang Y, Yang T, Wu C. Analysis on mal-operation of UR T35/T60 transformer differential protection [J]. Power Syst Prot Control. 2009;37(14):118–21. (In Chinese).
6. Lv L, Luo J, Huang Z, Ou Z. Simulation and analysis of the inrush of power transformer [J]. Relay. 2007;35(5):4–6. (In Chinese).
7. Zong HL, Jin HF, Zhu Z, Zhang S. Transformer inrush detected by the variation of magnetizing impedance [J]. Proc CSEE. 2001;21(7):91–4. (In Chinese).
8. Li Y. The emulation and analysis of the inrush and malfunction current of power transformer [J]. Relay. 2002;30(6):29–32. (In Chinese).
9. Wang X. The emulation and analysis of the inrush and malfunction current of power transformer [J]. High Volt Electr Appar. 2003;39(6):11–4. (In Chinese).
10. He X, Shu H, Li L. Analysis and countermeasure of sympathetic inrush in operating transformers [J]. Electr Appl. 2006;25(4):43–6. (In Chinese).
11. Zhang X, He B, Zhang J. Principle and influencing factors of the transformer sympathetic inrush [J]. Autom Electr Power Syst. 2005;29(6):15–9. (in chinese).

Chapter 87
Influence of the Snubber Circuit on the Commutation Process of High-Power Five-Level Converter

Shuai Dong, Chongjian Li, Chunyi Zhu and Chengsheng Wang

Abstract The snubber circuit is used to suppress the internal overvoltage and overcurrent of the electronic devices and reduce the switching losses of the device. This chapter studies the commutation process of the high-power five-level bridge converter, especially the influence of the snubber circuit on the commutation process. Each parameter of the snubber circuit has been verified to have an influence on the changes of the integrated gate-commutated thyristors' (IGCT) turn-on and turnoff voltage by simulation. This can provide an important theoretical basis and guidance for the design and control of the practical high-power multilevel system. And the parameters of the snubber circuit designed by this way have been successfully applied to the inverter system of 20 MW.

Keywords Snubber circuit · Five level · Commutation process · Converter with IGCTs

87.1 Introduction

The snubber circuit is used to suppress the overvoltage and overcurrent of power devices and absorb the energy of switching process. Usually, the parameter of the snubber circuit is extrapolated by experience in engineering and simulation [1–2]. This chapter studies the high-power five-level neutral-point-clamped (NPC)/H-

S. Dong (✉) · C. Li · C. Zhu · C. Wang
Automation Research and Design Institute of Metallurgical Industry,
100071 Beijing, China
e-mail: shuai22134307@163.com

C. Li
e-mail: licj_arim@163.com

C. Zhu
e-mail: zcyi955@gmail.com

C. Wang
e-mail: wangcs@aritime.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_87

bridge inverter and analyzes the influence of commutation process on the snubber circuit so as to provide an important theoretical basis and guidance for the design and control of the practical high-power multilevel system.

87.2 Structure of Five-Level Inverter

Currently, there are two kinds of multilevel inverter topologies: the multilevel capacitive suspension inverter and the NPC/H-bridge inverter [3]. This chapter studies the five-level NPC/H bridge inverter with its single-phase topology, which is shown in Fig. 87.1.

87.3 Analysis of the Snubber Circuit in Commutation Process

The snubber circuit is used to suppress the internal overvoltage and overcurrent of the electronic devices. In this part, the running way of the snubber circuit in commutation process is analyzed clearly [4–6].

Each phase of five-level NPC/H-bridge inverter is composed of two NPC legs in H-bridge form. In order to analyze the commutation process, the discrete output status of NPC leg is represented in Table 87.1.

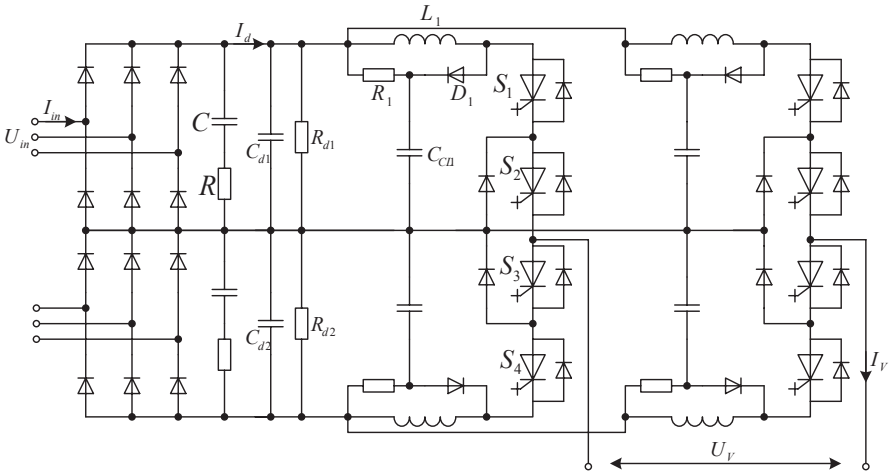


Fig. 87.1 Single-phase topology of five-level NPC/H-bridge inverter

Table 87.1 Status of the single arm of the NPC/H-bridge inverter

Number	Switch status	S1	S2	S3	S4	The output voltage
P	1100	Open	Open	Stop	Stop	$U_d/2$
O	0110	Stop	Open	Open	Stop	0
N	0011	Stop	Stop	Open	Open	$-U_d/2$

NPC neutral point clamped

87.4 Simulation and Experiment

Considering the miscellaneous inductor parameters, the test circuit of a single IGCT can be set up as shown in Fig. 87.2. The awkward recovery character of the fast recover diode is simply simulated by the RC circuit parallel to diodes D_s1 and V_f. The di/dt snubber circuit consists of inductor (L_s1), resistor (R_s1), diode (D_s1), and capacitor (C_s1). Lss1 is the miscellaneous inductor in the DC side capacitor branches. Lss2 is the miscellaneous inductor in the capacitor branch of the snubber circuit. Lss3 is the miscellaneous inductor in the resistor branch of the snubber circuit. Lss4 is the miscellaneous inductor in the diode branch of the snubber circuit. Lss5 is the miscellaneous inductor in the branch of IGCT.

Regardless of the miscellaneous inductor, the parameter of the snubber circuit is shown below [7, 8]:

$$\begin{cases} L_{s1} \geq V_{\mathrm{DC}} / \left(\mathrm{d}i / \mathrm{d}t\right) \\ C_{s1} \geq L_{s1} D^4 I_{s1} / \left(K \cdot \Delta V\right) \\ R_{s1} \leq L_{s1} / C_{s1} / \sqrt{2D} \\ \Delta V = V_{\mathrm{DRM}} - V_{\mathrm{DCMAX}} \end{cases}, \tag{87.1}$$

where di/dt is the turnoff capacity of the free-wheel diode parallel to IGCT, V_{DRM} is the peak voltage IGCT can stand, V_{DCMAX} is the maximum voltage of DC side capacitor, and I_{s1} is the maximum current passing L_s1.

Suppose the turnoff capacity of the free-wheel diode parallel to IGCT is 1000 A/μs, $V_{\mathrm{DCMAX}} = 2800\,\mathrm{V}$, $V_{\mathrm{DRM}} = 4500\,\mathrm{V}$, $I_{s1} = 2600\,\mathrm{A}$, D is about 0.8, K is about 0.9, and the result from formula (1) is $L_{s1} \geq 4.3\ \mu\mathrm{H}$. Now assign $L_{s1} = 5\ \mu$ERROR TRANSLATION H, it can be figured out that $C_{s1} \geq 3.5\ \mu$ERROR TRANSLATION F, $R_{s1} \leq 0.75\Omega$; the typical value as $C_{s1} = 5\ \mu\mathrm{F}$, $R_{s1} = 0.5\Omega$; and assign the value of miscellaneous inductor as the typical value, thus Lss1 = Lss2 = Lss3 = Lss4 = 0.1 μH, Lss5 = 0.3 μH. As to the effect of all the elements in the snubber circuit on IGCT turnoff character, if the effect of some element is obtained, then all the others should be the typical value. The effect of elements in the snubber circuit on IGCT turnoff character is shown in Figs. 87.3, 87.4, and 87.5.

Based on the result of simulation shown in Figs. 87.3, 87.4, and 87.5:

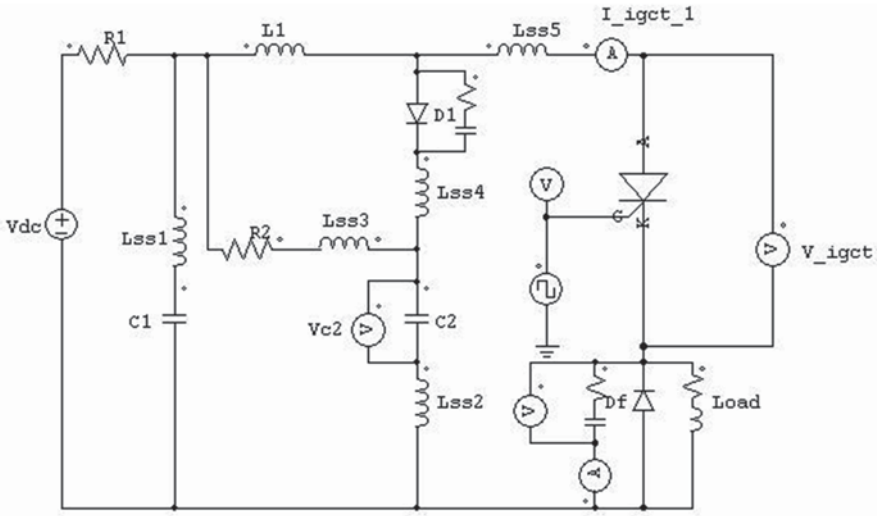


Fig. 87.2 Simulation of one IGCT test circuit. *IGCT* integrated gate-commutated thyristor

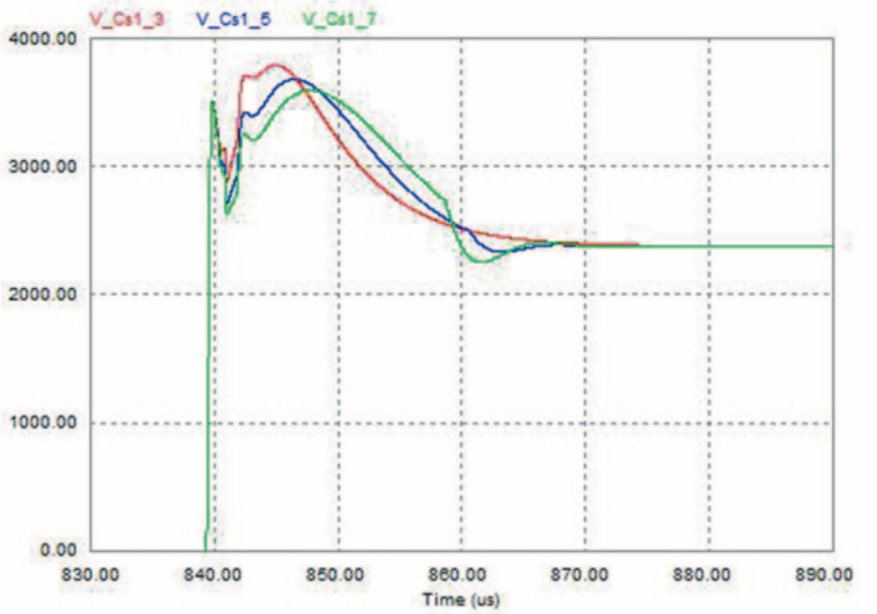


Fig. 87.3 Effect of electric capacity

1. The value of capacitor has an effect on the peak value of IGCT turnoff voltage wave shape. When the value of capacitor increases, the peak value of IGCT decreases.
2. The value of capacitor has a little effect on the absorb time.
3. The value of capacitor has an effect on the transition time between the two peaks. When the value of capacitor increases, the transition time between the two peaks decreases.

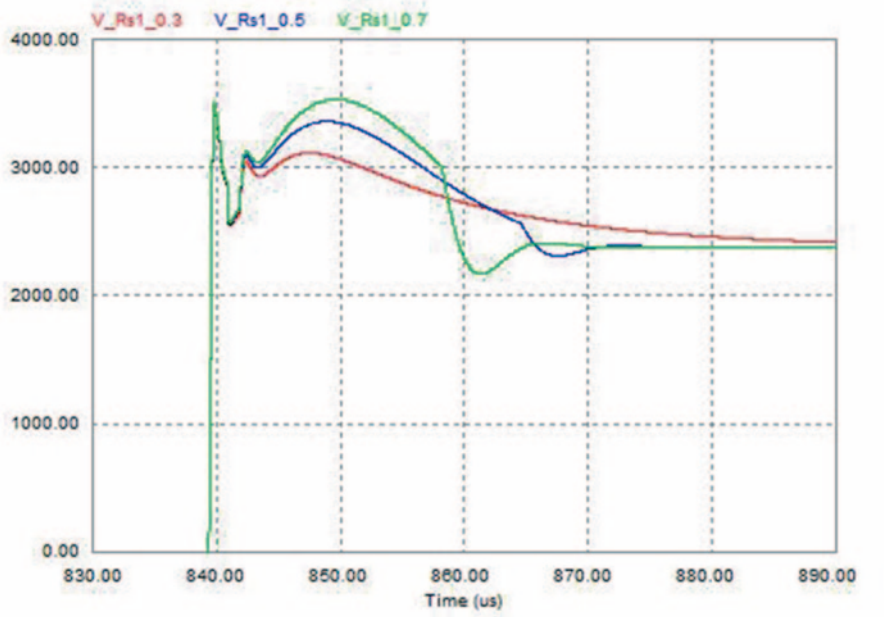


Fig. 87.4 Effect of resistance

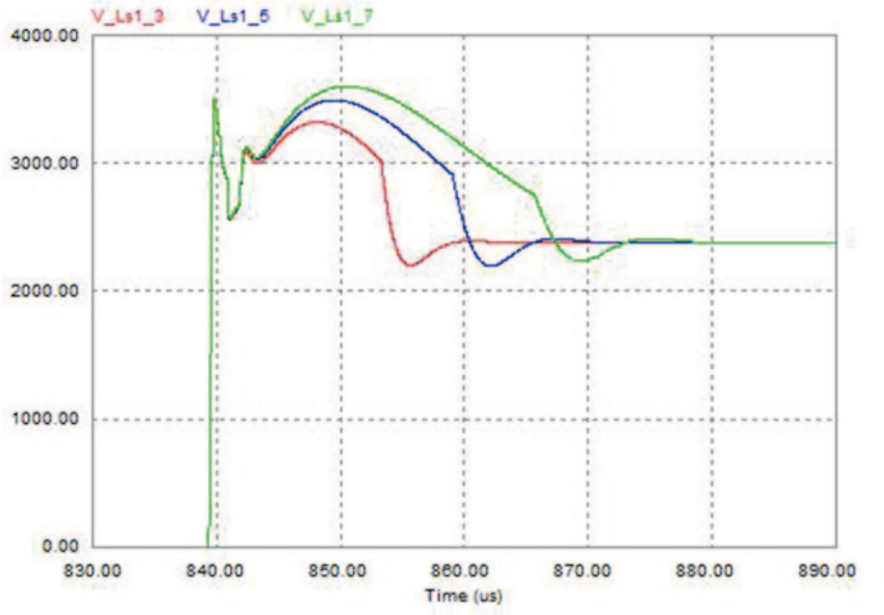


Fig. 87.5 Effect of inductance

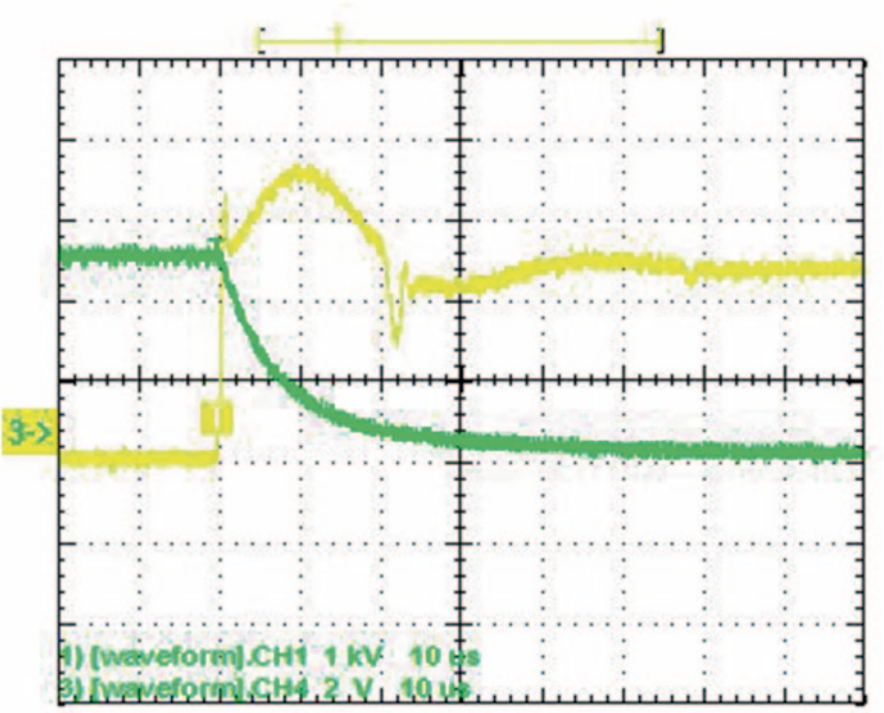


Fig. 87.6 Shutdown voltage waveform of the IGCT five-level inverter (the DC voltage is 5000 V and the peak load current is 2500 A)

4. The value of resistor has an effect on the second voltage peak. When the value of resistor increases, the peak value of IGCT decreases.
5. The value of resistor has an effect on the transition time to steady state. When the value of resistor increases, the transition time to steady state decreases.
6. The value of inductance has an effect on the second voltage peak. When the value of inductance increases, the peak value of IGCT increases.
7. The value of inductance has an effect on the transition time to steady state. When the value of inductance increases, the transition time to steady state increases.

In practical applications, we usually use formula (1) to calculate the parameter of the snubber circuit. Firstly, the value of inductance should be lower to reduce the volume of the electric circuit and the second voltage peak, and the transition time to steady state at meanwhile; secondly, the value of capacitor should be lower and the value of resistor should be higher. This can ensure the IGCT shutoff withstand voltage does not exceed a safe range. Finally, we can use the low-sense bus bar and reasonable structure layout to reduce the connection inductance; besides, we can also use no inductance components in the capacitance DC side and snubber circuit to reduce the stray inductance of element itself. It can effectively reduce the stray inductance of the circuit so as to ensure the IGCT shutoff withstand voltage does not exceed a safe range (87.6).

87.5 Conclusion

This chapter studies the commutation process of the high-power five-level NPC/H-bridge converter, especially the influence of the snubber circuit on the commutation process. Each parameter of the snubber circuit has been verified to have the influence on the changes of the IGCT turn-on and turnoff voltage by simulation. This can provide an important theoretical basis and guidance for the design and control of the practical high-power multilevel system. The results of this chapter have been used to guide the establishment of a five-level NPC/H-bridge inverter with IGCTs of 20 MW. It can run in good condition and the turnoff voltage spikes of IGCTs are in the safe range. This can prove the correctness of the results of this chapter.

References

1. Steimer PK, Grüning HE, Werninger J. IGCT-a new emerging technology for high power, low cost inverter. IEEE Industry Applications Society Annual Meeting New Orleans, ABB Review, Louisiana; 1998. p. 34–42.
2. Eicher S, Bernet S, Steimer P, Weber A. The 10 kV IGCT—a new device for medium voltage drives. Industry Applications Conference, 2000. Conference Record of the 2000 IEEE (Vol. 5), Rome; 2000. p. 2859–65.
3. Tong Y, Zhang C, Xie L. Drive technology for 4000A/4500V series IGCT. Trans Chin Electr Soc. 2010;25(8):110–5. (in Chinese).
4. Wang C, Li C, Li Y. Research on three-level AC-DC-AC converter equipped with IGCTs. Trans Chin Electr Soc. 2007;22(8):24–7. (in Chinese).
5. Lan Z, Li C, Zhu C. Large power high voltage converter equipped with integrated gate commutated thyristors. J Zhejiang Univ Eng Sci. 2007;41(10):1674–8. (in Chinese).
6. Li C, Wang C, Zhu C. Research on the transient state of large power voltage source inverter with IGCTs. Asia—pacific international symposium on electromagnetic compatibility (AFEMC). Engineering Science, Beijing, China: China Electrotechnical Society; 2010. p. 218–21.
7. Zhao Z, Zhang H, Yuan L. Failure mechanism and protection strategy of high voltage three-level inverter based on IGCT. Trans Chin Electr Soc. 2006;2 1(5):1–6. (in Chinese).
8. Yuan L, Zhao Z, Liu J. Design simulation and analysis of IGCT-based high power NPC inverter. Electr Power Autom Equip. 2004;24(1):46–9. (in Chinese).

Chapter 88
Design of a Low-Power Humidity and Temperature Collector Based on STM32

Zhihao Liu, Jianli Shi and Wenhui Peng

Abstract STM32 is a microcontroller based on CortexTM-M3 core; it plays an important role in the embedded control field. This chapter introduces a low-power humidity and temperature collector with SHT10 humidity and temperature sensors. This chapter introduces and implements the communication time sequence of SHT10. A calibration text proves that the system can measure temperature and humidity correctly. It is demonstrated by experiments that this low-power collector can collect data for 30 days and satisfy the design index; besides, the principle of low power can be used in other embedded control field.

Keywords Temperature and humidity collector · Low-power mode · STM32F103

88.1 Introduction

STM32 is widely used in many control applications. It is a microcontroller based on CortexTM-M3 core, which sets new standards in capability and cost as well as low-powered operation and real-time control. The series mainly includes three subseries: "A" means high applications, "R" means real-time applications series, and "M" means cost-sensitive and microcontroller applications. STM32 belongs to the "M" series, which is designed for the coalescence of high capability and low power [1].

Most of the humidity and temperature collectors on the market do not have the ability of recording and low power at the same time; in addition, they are expensive. This chapter designs a humidity and temperature collector based on STM-32F103RET6. Its system can work for about 30 days with the ability of recording and a stop mode has been adopted for the purpose of power saving. The sensor is SHT10 humidity and temperature sensor. The collection period takes 5 min, and 18,650 lithium battery works as the power supply. The data can be recorded in the secure digital (SD) card. The collector can support high-speed universal serial bus

Z. Liu (✉) · J. Shi · W. Peng
Naval Submarine Academy, 266041 Qingdao, China
e-mail: 396198831@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_88

(USB) communication and serial ports communication. The structure is from the practical design and the rest of this chapter is organized as follows.

88.2 Design of Collector

88.2.1 Hardware Design

The collector in this chapter is shown in Fig. 88.1 with the proposed structure diagram shown in Fig. 88.2. It mainly includes the power supply module, the high-speed USB communication module, STM32 IC, SD card recording module, the humidity and temperature sensor module, the serial ports communication module, and joint test action group (JTAG) debug online module.

STM32 gets power from ADP339-A3.3 voltage regulator. It is woken up every 5 min by real-time clock (RTC) to collect data received from SHT10 sensor by user-defined communication protocol. The data on humidity and temperature data are written into SD card in the form of hexadecimal [2].

88.2.2 Configuration of Communication

The communication sequence of SHT10 is shown in Fig. 88.3.

At the first step, the sensor is powered up and the slew rate in the course cannot be below 1 V/ms. The sensor needs 11 ms to get to the sleep state after power-up. No command must be sent before that time. To initiate a transmission, we should issue a transmission start sequence which consists of a lowering of the *data* line while serial clock (SCK) is high, followed by a low pulse on SCK, and raising *data* again while SCK is still high. The subsequent command consists of three address bits and five command bits. The sensor indicates proper reception of a command by pulling the *data* pin low after the falling edge of the eighth SCK.

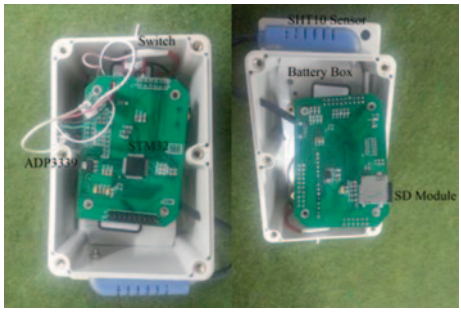


Fig. 88.1 Photograph of system prototype

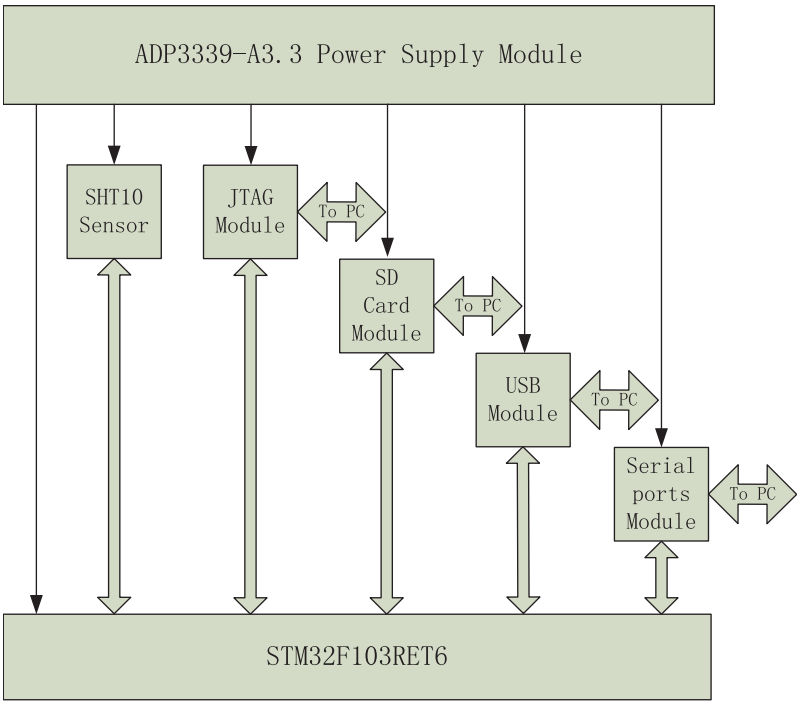


Fig. 88.2 Structure diagram of collector

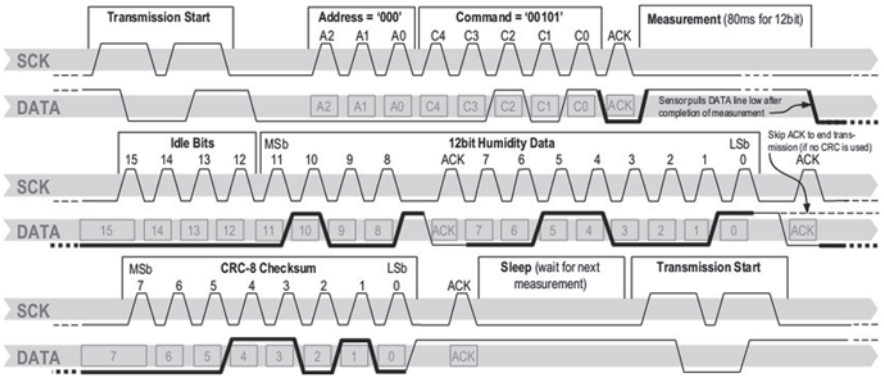


Fig. 88.3 Communication sequence of SHT10 sensor

The *data* line is released after the falling edge of ninth SCK. The list of SHT10 commands is shown in Table 88.1.

After issuing a measurement command ("00000101" for relative humidity, "00000011" for temperature), the controller has to wait for the measurement to complete. This takes a maximum of 20/80/320 ms for an 8/12/14-bit measurement.

Table 88.1 List of commands

Command	Code
Reserved	0000×
Measure temperature	00011
Measure relative humidity	00101
Read status register	00111
Write status register	00110
Reserved	0101×–1110×
Soft reset	11110

To sign the completion of measurement, SHT10 pulls *data* line low and enters the idle mode. The controller must wait for this *data* ready signal before restarting SCK to readout the data. Then the measurement data are stored after readout. The device automatically returns to the sleep mode upon completion of measurement and communication.

The sensor connects to the microcontroller through PIN6 and PIN7 of general purpose input/output port B (GPIOB) in STM32. PIN6 is the clock communication port and PIN7 is the data communication port. The output mode of PIN6 pushes the pull output mode and the speed of transmission is 50 MHz. PIN7 pushes the pull output/input with the speed of 50 MHz [3].

The Transmission Start sequence is implemented by writing SCK and *data* line in SHT10. After the configuration of PIN7, a series of codes should be operated. Pulling SCK is low; pulling SCK is high; pulling *data* is low; pulling SCK is low; pulling SCK is high; pulling *data* is high; pulling SCK is low. There should be time delay between each operation. After transmission start sequence is completed, the sensor will collect and write temperature or humidity data by 0×03 or 0×05.

88.2.3 Error Correction

In order to improve the measurement precision, once STM32 is woken up, it will collect temperature and humidity data ten times. Then the average value of these data will be considered as the reasonable environment parameter. It is the average value that will be written in SD card.

A test has already been carried out to calibrate the error of sensor. A standardized temperature and humidity collector and a proof box are essential, in which the proof box supplies the fixed parameter of environment. The standardized collector supplies accurate temperature and humidity data. Accurate data can calibrate the collector in this chapter. The error of collector is shown in Fig. 88.4. It is reasonable to get the conclusion that the measurement precision can satisfy the requirement.

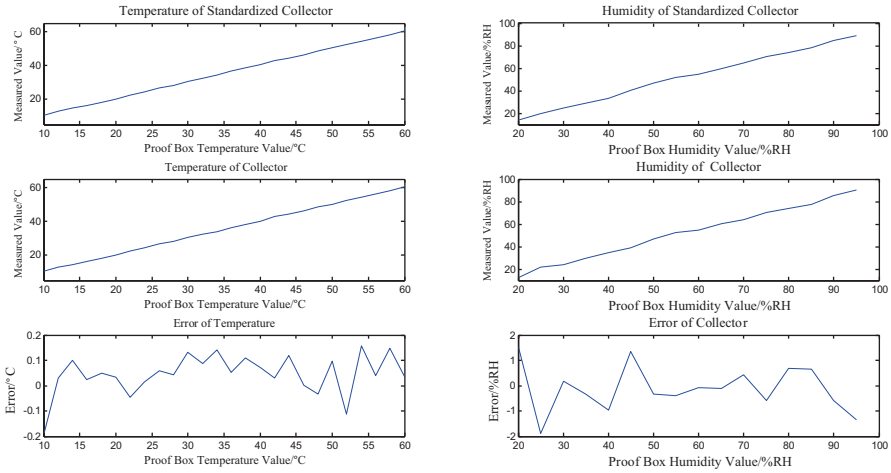


Fig. 88.4 Calibration of collector

88.3 Clock System and Power Test

88.3.1 Clock System of STM32

The clock system is the precondition to make STM32 and its peripherals running properly. What is more, the frequency of system clock and its peripherals has great influence on power. STM32 has internal RC (Resistance and capacitance) oscillator, the oscillator can supply internal phase-locked loop (PLL) clock. This PLL clock can achieve the highest frequency requirement of microcontroller, that is, 72 MHz; but the internal oscillator is inaccurate and unsteady compared with the external clock. In this sense, the collector chooses the external clock as alleged high-speed external oscillator (HSE). This collector system chooses an oscillator whose oscillating frequency is 8 MHz [4].

The external oscillator will generate 8-MHz clock signal. This signal will be 72 MHz after PLL, which will be nine octaves. The system will run with this frequency. The clock frequency of peripherals on APB1 is half of the system clock and the clock frequency of peripherals on APB2 is the system clock. APB1 supports peripherals whose clock frequency is not very high. Its maximum frequency is 36 MHz. APB2 supports peripherals whose clock frequency is higher than 36 MHz. Its maximum frequency is 72 MHz. Almost all peripherals choose APB1 and APB2 for their clock.

STM32 has two power domains. One is the domain of main system and external devices. The other is the backup domain, which includes ten 16-bit registers, RTC, and independent watchdog. RTC and watchdog can keep running to wake up the main system or reset in the low-power mode. This collector uses external low-speed

clock as RTC and wakes up main power domain by alarm interrupt. The stored value in alarm register is 32,767 [5]. With the consideration of volume, it does not design the backup battery in this collector. RTC gets power from the output of ADP3339-A3.3 module directly.

The clock configuration of the collector in this chapter is shown in Fig. 89.5.

88.3.2 Power Test

The power supply diagram of STM32 chip is shown in Fig. 88.6. The collector's power comes from a 4.2-V lithium battery with the capacity of about 2600 mAh. ADP3339-A3.3 regulator can fix the output voltage in 3.3 V to supply power for chip and peripherals [6]. The power supply module of collector is shown in Fig. 88.7. The transfer efficiency of ADP3339-A3.3 is about 90 %. Considering the efficiency, current, and voltage, the available battery capacity is about 3000 mAh. When the system clock is 72 MHz, the current of the system is about 100 mA measured by the multimeter. A test has been carried out in which the period takes 5 min and the battery is full. STM32 writes data in the SD card through secure digital input output (SDIO) interface protocol every 5 min. The test continues for about 26 h before the battery is empty. We can conclude that the system without low-power mode will cost much power. The collector without low-power mode could not meet the needs of long work time in closed environment.

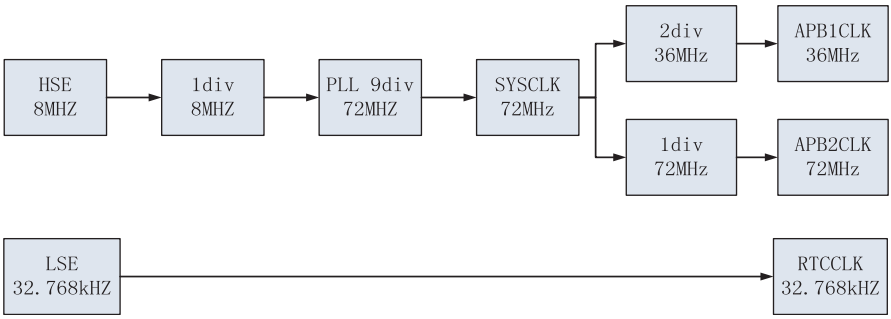


Fig. 88.5 Clock configuration of collector

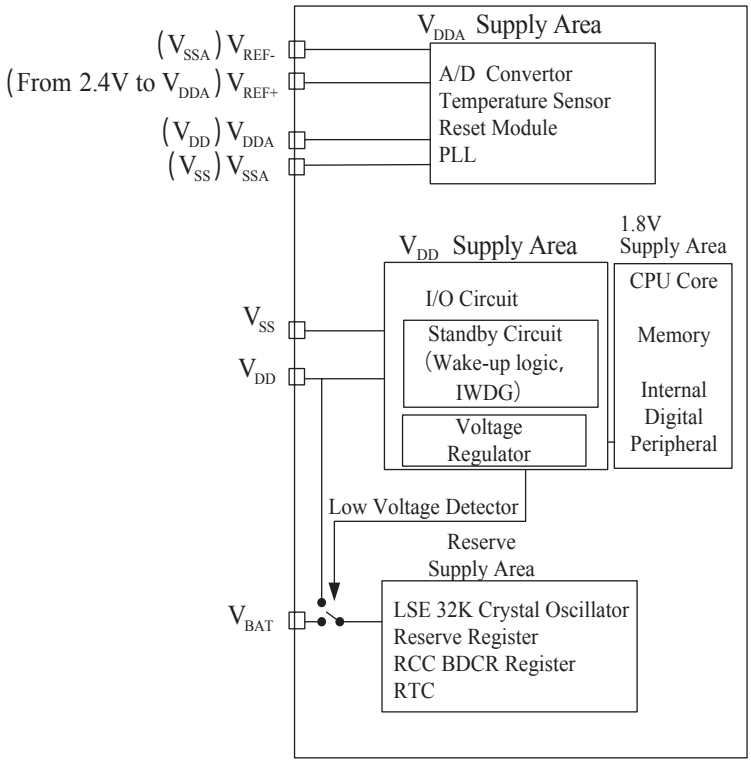


Fig. 88.6 Power supply diagram of STM32 chip

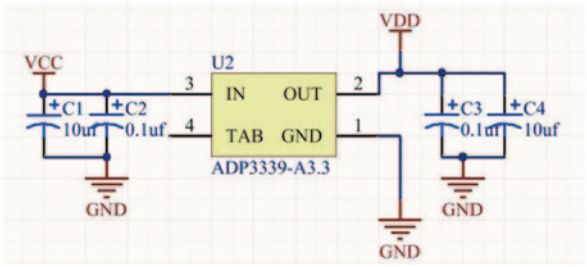


Fig. 88.7 Power supply module of collector

88.4 Implementation of Low-Power Mode

88.4.1 Software Implementation of Low Power

STM32 has three kinds of low-power modes: the sleep mode, the stop mode, and the standby mode, in which, the stop mode is adopted in this collector. It is the key point of low-power mode. The stop mode achieves the lowest power consumption while retaining the content of static random-access memory (SRAM) and register. All clocks in the 1.8-V domain are stopped; the PLL, the high speed interal RC oscillator (HSI RC), and the HSE crystal oscillators are disabled. The voltage regulator can also be put either in normal or in low-power mode. In this mode, any external interrupt will wake up the central processing unit (CPU) [7, 8].

We must be sure to enable the power control clock and the RTC back up register, then enable the external 32.768-KHz low-speed oscillator and RTC after the socillator is steady and configured. Any operation must wait for the last operation. RTOFF state bit can judge whether RTC register is updated. The writing is permissible when the RTOFF state bit is "1." The configuration of RTC register is shown in Fig. 88.8.

88.4.2 Contrast Test and Analysis

The collector without low-power mode collects data in the open air. The power comes from a 2600-mAh lithium battery and the battery is full. The output current from battery is 103 mA. The collection period takes 5 min. The collector works for 26 h and 20 min with 316 data points collected, with the collected data shown in Fig. 88.9.

At the opposite position are the same collector and the same test condition, but this collector has low-power mode. The output current from battery is 4 mA when it is in low-power mode and is 54 mA when it is woken up. Ten days after the beginning of the test, the voltage of the lithium battery is 3.96 V and the collector is still working under this condition. The number of collected data is 2936, which means the collector has been already working for 10 days and 4 h and 40 min. The collected data are shown in Fig. 88.10.

The battery voltage drops to 0.3 V in 10 days. In the case of calculation according to this speed, it should be 30 days until the voltage drops to 3.35 V, which will meet the needs for long work time in a closed environment.

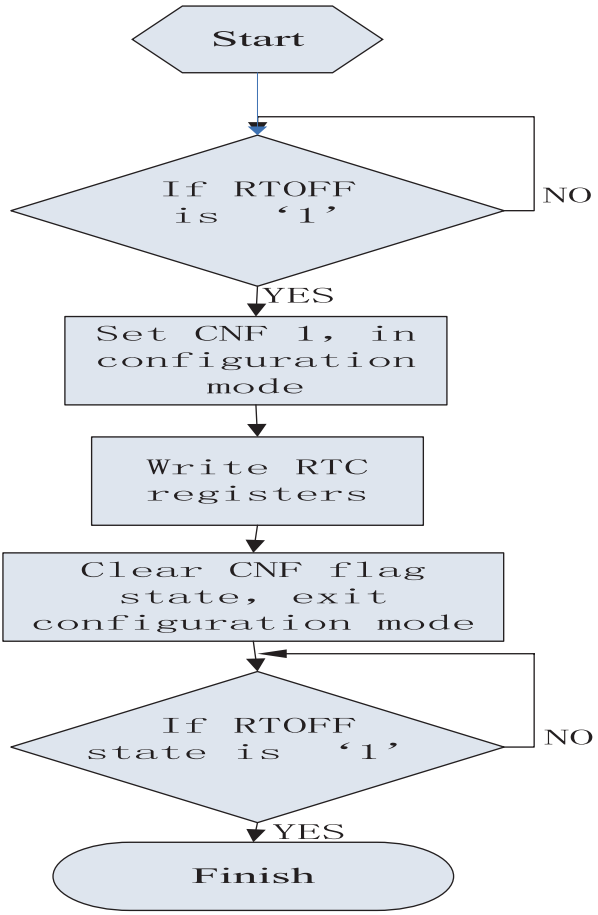


Fig. 88.8 Configuration of RTC register

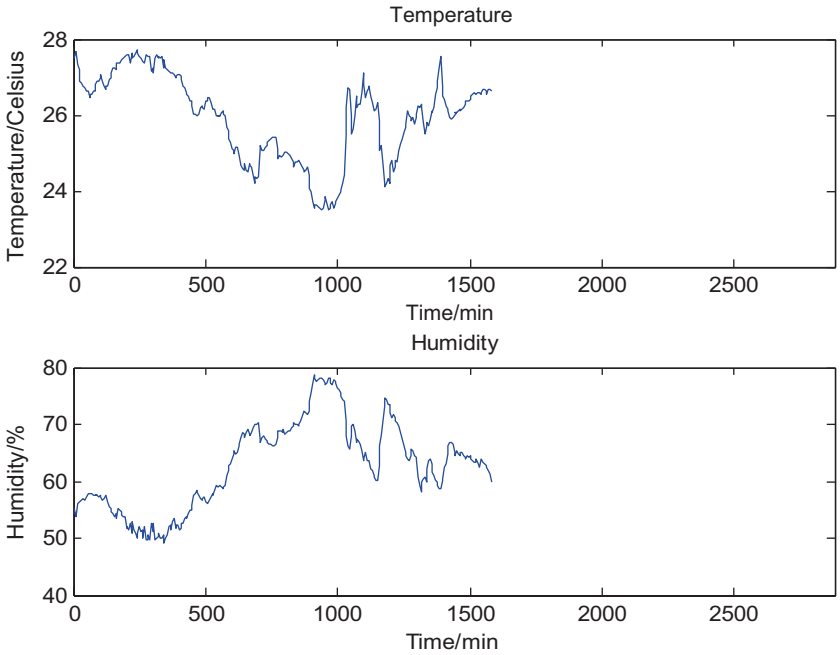


Fig. 88.9 Collected data without low-power mode

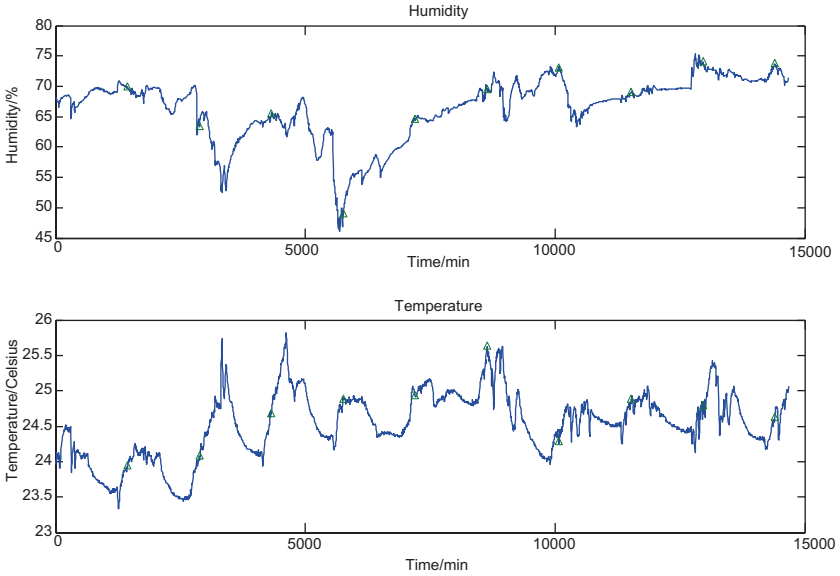


Fig. 88.10 Collected data with low-power mode

88.5 Conclusion

In this chapter, a type of temperature and humidity collector with recording function and low-power consumption is designed to meet the need of long time work in closed environment. Upon test, the work time is found to be 26 h. Under the contrast test, the collector with low-power mode can work about 30 days. The communication sequence of SHT10 sensor has already been explained and the configuration of the transmission start sequence has already been implemented. These methods can be used in other STM32F103x systems.

References

1. Zhijia C, Shuying S. Design of air tracking servo system based on STM32F103. Instrumentation Measurement, Computer Communication and Control, 2011 First International Conference on, IEEE; 2011. p. 891–4.
2. Yingchun K, Yue S. A Greenhouse Temperature and Humidity Controller Based on MIMO Fuzzy System. Intelligent System Design and Engineering Application (ISDEA), 2010 International Conference on, IEEE; 2010. p. 35–9.
3. Zhitao G, Jinli Y. Design of ocean intelligent sensor based on STM32. Test and Measurement, 2009. ICTM'09. International Conference on, IEEE; 2009. p. 124–7.
4. Tan N, Eriksson S. Low-power chip-to-chip communication circuits. Electron Lett. 1994;30(21):1732–3.
5. Ismail SM, Rahman ABMS, Islam FT. Low power design of Johnson Counter using clock gating. Computer and Information Technology (ICCIT), 2012 15th International Conference on, IEEE; 2012. p. 510–7.
6. Zhang R, Xiong G, Cheng C, et al. Control system design for two-wheel self-balanced robot based on the stepper motor. Service Operations and Logistics, and Informatics (SOLI), 2013 IEEE International Conference on, IEEE; 2013. p. 241–4.
7. Analog Devices, Inc. (2001–2011). ADP3339 DATASHEET [online]. www.analog.com. [4/6/2014].
8. Hungche C, Chao PC-P, Weichu L, et al. A new gas sensor of a thin-film diode and low-power, area-efficient readout circuit. Sensors, 2013 IEEE, IEEE; 2013. p. 1–4.

Chapter 89
Implementation Method for Class-J Power Amplifiers

Tao Chen, Xiaohong Sun, Feng Wang and Jianhui Wu

Abstract This chapter presents an implementation method for Class-J power amplifiers. The requirement for the load impedances of Class-J power amplifiers is analyzed first. Then an output matching structure is proposed for low-optimum-load impedance condition. It uses a short-terminated transmission line to supply the collector voltage of the power amplifier and meet the requirement for the second harmonic impedance. To verify this method, a one-stage common emitter power amplifier using 2-μm gallium arsenide heterojunction bipolar transistor (GaAs HBT) process has been designed and a Class-J power amplifier is implemented with the component on printed circuit board (PCB). The implemented Class-J power amplifier exhibits a 55% collector efficiency with the saturated output power of 34.2 dBm at the frequency of 2.14 GHz and 3.5 V supply voltage.

Keywords Power amplifier (PA) · Class J · Efficiency

89.1 Introduction

The power amplifier is usually the largest energy consumption component in the wireless communication systems. There are two main methods in respect of the design of high-efficiency power amplifiers with one based on the switching mode and the other based on harmonic tuning (HT) strategies. The power amplifier working in switching mode provides high efficiency but poor linearity [1]. One of the harmonic tuning techniques named Class-G design strategy was proposed and implemented by using the second harmonic manipulation of the output voltage [2]. Another second harmonic manipulation technique called Class-J design was proposed by Steve C. Cripps [3]. The difference between these two designs was analyzed

T. Chen (✉) · X. Sun · J. Wu
National ASIC System Engineering Research Center, Southeast University, 210096 Nanjing, China
e-mail: seu_ct@sina.com

F. Wang
Suzhou Innotion Tech. Co., Ltd., 215123 Suzhou, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_89

by Elisa Cipriani [4]. Class-G power amplifiers provide higher output power and efficiency than Class-J power amplifiers at the expense of bandwidth. Class-G power amplifiers need a strict design of the input second harmonic voltage to get a proper phase of the collectors' second harmonic current. It is easy to realize the design procedure of Class-J power amplifiers; and the nonlinear output capacitor of transistors would enhance the output power and efficiency of Class-J power amplifiers [5]. This chapter designs a Class-J power amplifier by using a short-terminated transmission line to provide a low value of second harmonic load impedance.

In Sect. 89.2, the mathematical statements of Class J are reviewed. The Class-J power amplifier is designed by using a short-terminated transmission line to reach a proper second harmonic impedance in Sect. 89.3. Finally, the simulation and measurement results are presented in Sect. 89.4.

89.2 Theory of Class-J Power Amplifier

The device, operated in Class-J mode and considered as a voltage-controlled current source, is always biased at Class-B or deep Class-AB condition; therefore, the collector current is a half-rectified sinusoidal waveform. The output voltage waveform depends on the fundamental and harmonic load impedance. The Class-J power amplifier makes use of the second harmonic voltage to peak the voltage waveform in the high-voltage region and flatten the voltage waveform in the low-voltage region, which would increase the maximum fundamental voltage component. Class-J power amplifier usually includes a capacitor as the second harmonic load, and the fundamental load is not a pure resistor, but with a reactive component.

For the sake of simplicity in illustration, the knee voltage is assumed to be zero, and the device is biased at Class-B condition. The collector current can be expressed as:

$$I_c(\theta) = \frac{I_{max}}{2}\left(\cos(\theta) + \left|\cos(\theta)\right|\right) \tag{89.1}$$

where I_{max} represents the maximum collector current. This equation could be expanded as Fourier series. When the third harmonic and above are neglected, the collector current is simplified as:

$$I_c(\theta) = I_{max}\left(1/\pi + 1/2\cos(\theta) + 2/(3\pi)\cos(2\theta)\right) \tag{89.2}$$

The collector voltage of Class-J mode can be expressed as [6]:

$$V_J(\theta) = V_{DC}\left(1 - \sqrt{2}\cos\left(\theta + \frac{\pi}{4}\right) + 1/2\cos\left(2\left(\theta + \frac{\pi}{4}\right)\right)\right) \tag{89.3}$$

The fundamental and second harmonic load impedances can be calculated as:

$$Z_1 = R_{\text{opt}} + jR_{\text{opt}} \tag{89.4}$$

$$Z_2 = -\,j\frac{3\pi}{8}R_{\text{opt}} \tag{89.5}$$

where $R_{\text{opt}} = V_{\text{DC}}/(I_{\text{max}}/2)$, which is the optimum load in Class-B condition.

According to Eq. (89.4), the real part of the fundamental load impedance is equal to Class-B power amplifier, so the maximum output power is equal to Class-B condition. The prior analysis is based on the linear output capacitor, which would not change when the output power is increased; however, in real condition, the output capacitor is usually nonlinear, which could bring negative resistance in the second harmonic load impedance [5]. This negative second harmonic load resistor would shift the voltage phase and enable the real part of fundamental load impedance to become larger than that at Class-B condition, which would result in the increase of the maximum output power and efficiency.

89.3 Class-J Power Amplifier Design

A simple one-stage common emitter power amplifier is designed by using a 2-μm gallium arsenide heterojunction bipolar transistor (GaAs HBT) process. The power cell is biased by an active bias circuit. The chip does not include input and output matching circuits, which will be implemented on the printed circuit board (PCB).

A 2.14-GHz Class-J power amplifier is designed with this power amplifier chip. The supply voltage is 3.5 V, and the optimum load could be calculated as 1.5 Ω. According to Eq. (89.5), the reactive value of the second harmonic impedance is also very low. The low reactive value could be implemented by a high-value lumped capacitor; however, the available capacitors of high value are not as many as the lumped capacitors of low value. In this design, the collector voltage is supplied by a short-terminated transmission line, which is a little shorter than the quarter-wave; and this transmission line could also provide a low second harmonic capacitive reactance. The impedance of this short-terminated transmission line is:

$$Z_{\text{in}}(l) = jZ_0 \tan\left(\frac{2\pi}{\lambda}l\right) \tag{89.6}$$

where Z_0 represents the characteristic impedance of the transmission line, λ is the wavelength, and l represents the length of the transmission line.

The simulated impedance of transmission line is shown in Fig. 89.1. In the figure, the length of transmission line is normalized to $\lambda/4$. In the fundamental frequency, the load is a high inductive reactance; and in second harmonic frequency, the load is a low capacitive reactance. This transmission line would help reach the

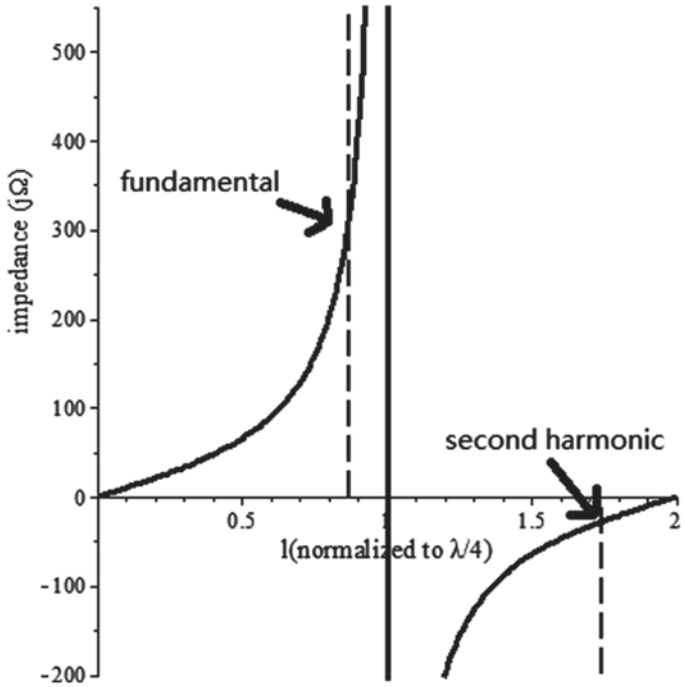


Fig. 89.1 Input impedance of short-terminated transmission line

required second harmonic load impedance without affecting the fundamental load matching circuit.

As affected by the transistors' nonlinear output capacitor and bonding wire, the load impedances are required to de-embed these influences. The input matching is used to obtain good input reflection feature. The designed circuit is shown in Fig. 89.2, where the output matching network is used to transfer the load to the optimum impedance in fundamental frequency, and the short-terminated trans-

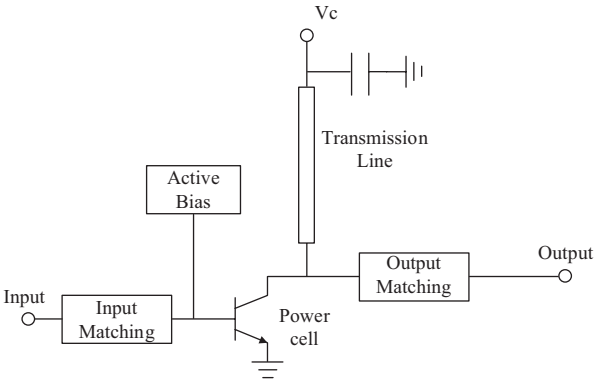


Fig. 89.2 Designed Class-J power amplifier

mission line is used to get the optimum second harmonic load impedance for the amplifier.

89.4 Measurement Results

The implemented PCB board is shown in Fig. 89.3. The output matching circuit has been designed by using a distributed network and lumped capacitors. The supply voltage is 3.5 V. The transmission line bias structure limits the bandwidth of the amplifier. The simulated and measured results of small signal scattering parameters are shown in Fig. 89.4. Figure 89.5 shows the simulated and measured results of power gain and output power at the frequency of 2.14 GHz. The saturated output power could reach 34.2 dBm. The simulated and measured results of the power efficiency versus output power are plotted in Fig. 89.6. A peak efficiency of 55 % is reached at the frequency of 2.14 GHz with the output power of 34.2 dBm.

Table 89.1 lists the comparison of recent studies reported in the literatures and this work. Except this work, all the other studies use gallium nitride (GaN) high-electron-mobility transistor (HEMT) to design Class-J power amplifier; and all the supply voltages are higher than this work, so the output matching is easier to implement. This work is the first design of high output power Class-J amplifier under the supply of low voltage.

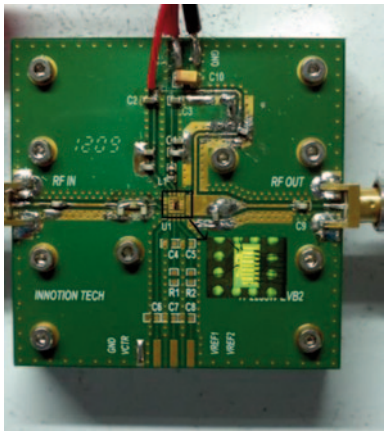


Fig. 89.3 Test PCB board. PCB printed circuit board

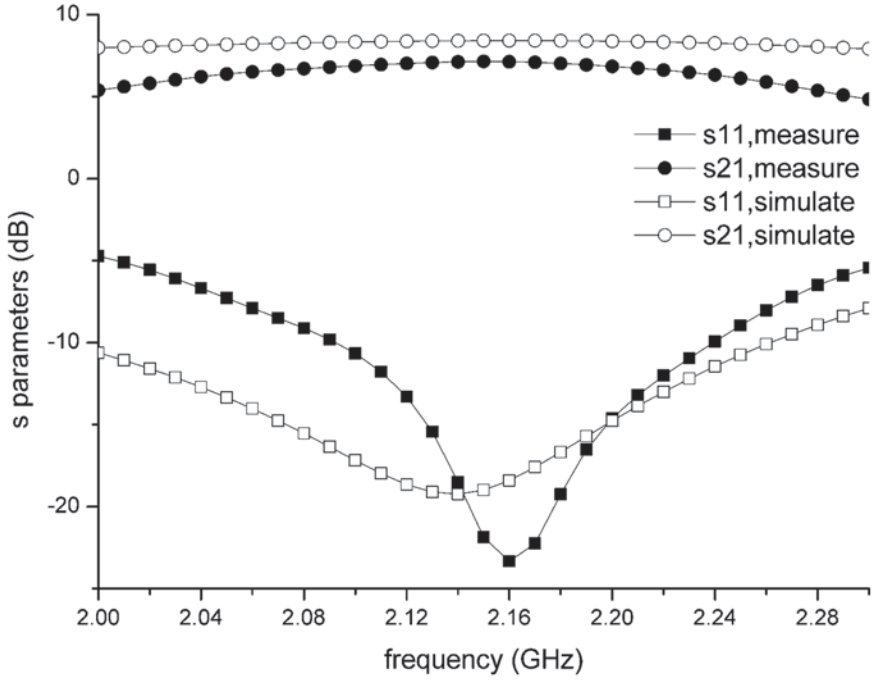


Fig. 89.4 Scattering parameters

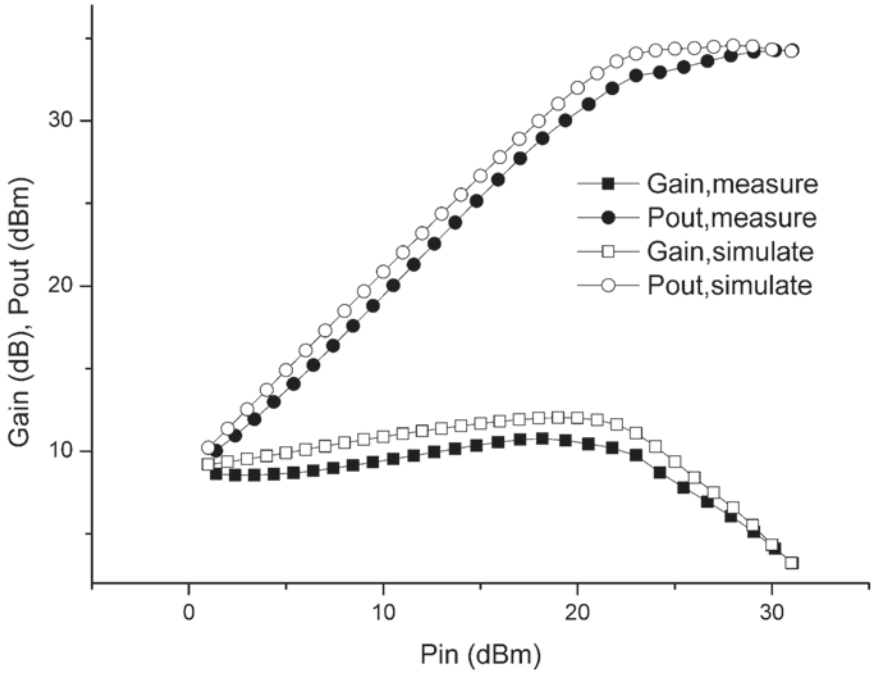


Fig. 89.5 Gain and output power as functions of input power

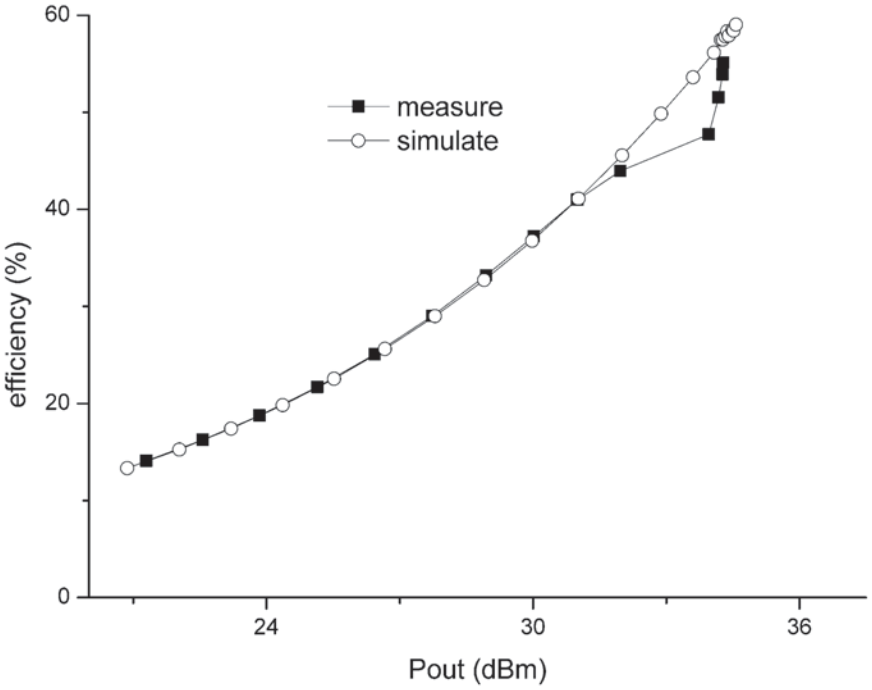


Fig. 89.6 Collector efficiency versus output power

Table 89.1 Comparison of Class-J PA reported in the literatures and this work

Ref	Year	Frequency (GHz)	Efficiency (%)	P_{out} (dBm)	Bias voltage (V)	Device process
[7]	2011	2.08–2.2	60	39	28	GaN HEMT
[8]	2011	2.3–2.7	60	40	45	GaN HEMT
[4]	2012	2.6–3.3	46	36.8	25	GaN HEMT
[9]	2012	1.7–2.24	62	40	28	GaN HEMT
[10]	2013	2.14	50–58	24–27	15	GaN HEMT
This work		2.14	55	34.2	3.5	GaAs HBT

GaAs gallium arsenide, *GaN* gallium nitride, *HBT* heterojunction bipolar transistor, *HEMT* high-electron-mobility transistor, *PA* power amplifier

89.5 Conclusion

In this chapter, a Class-J power amplifier is implemented. It is biased by a short-terminated transmission line. This structure could tune the second harmonic load impedance without the influence on fundamental load impedance, and it makes it easy for the Class-J matching requirements to fulfill the low optimum load impedance condition; however, the bandwidth of this amplifier is also limited. The power amplifier demonstrates 55 % of the collector's efficiency with the saturated output power of 34.2 dBm at the frequency of 2.14 GHz and 3.5 V supply voltage.

References

1. Sokal NO, Sokal AD. Class E—a new class of high efficiency tuned single ended switching power amplifiers [J]. IEEE J Solid-State Circuits. 1975;10(3):168–76.
2. Colantonio P, Giannini F, Leuzzi G, Limiti E. Class G approach for low-voltage, high-efficiency PA design [J]. Int J RF Microw Computer-Aided Eng. 2000;10(6):366–78.
3. Cripps SC. RF power amplifiers for wireless communications [M]. 2nd ed. Boston: Artech House Publishers; 2006. pp. 68–77.
4. Cipriani E, Colantonio P, Giannini F. The effect of 2nd harmonic control on power amplifiers performances [C]. Microwave Integrated Circuits Conference (EuMIC), 2012 7th European. Amsterdam: IEEE; 2012. pp. 345–8.
5. Moon J, Kim J, Kim B. Investigation of a Class-J power amplifier with a nonlinear C_{out} for optimized operation [J]. IEEE Trans Microw Theory Tech. 2010;58(11):2800–11.
6. Cripps SC, Tasker P, Clarke A, Lees J, Benedikt J. On the continuity of high efficiency modes in linear RF power amplifiers [J]. IEEE Microw Wirel Compon Lett. 2009;19(10):665–7.
7. Mimis K, Morris KA, McGeehan JP. A 2 GHz GaN class-J power amplifier for base station applications [C]. Power Amplifiers for Wireless and Radio Applications (PAWR), 2011 IEEE topical conference on Phoenix: IEEE; 2011. pp. 5–8.
8. Tuffy N, Zhu A, Brazil TJ. Class-J RF power amplifier with wideband harmonic suppression [C]. Microwave Symposium Digest (MTT), 2011 IEEE MTT-S international. Baltimore: IEEE; 2011. pp. 1–4.
9. Guan Y, Chen WH, Feng ZH. High efficiency and wide band CLASS-J power amplifier using 2nd harmonic microstrip stub matching [C]. Microwave and Millimeter Wave Technology (ICMMT), 2012 international conference on. Shenzhen: IEEE; 2012. pp. 1–4.
10. Rezaei S, Belostotski L, Ghannouchi FM, Aflaki P. Integrated design of a class-J power amplifier [J]. IEEE Trans Microw Theory Tech. 2013;61(4):1639–48.

Chapter 90
Analysis of Circulating Currents in a Modular Multilevel Converter-Static Synchronous Compensator

Weiwei An, Yukun Sun and Liang Zhang

Abstract A new modulation method named CPPS is proposed in this chapter; besides, in order to identify the basic components and analytical expressions of arm current in a modular multilevel converter-based static synchronous compensator (MMC-STATCOM), an analysis is made with a mathematical method based on instantaneous power conservation under the modulation mode of CPPS. It proves the existence of the second circulating current and gives its analytical expressions as well as the relevant submodule (SM) capacitor voltage ripple. The analysis implies that the second circulation current cannot be eliminated by controllers. A three-phase MMC-STATCOM model with nine voltage levels applied to a 10-kV system is realized to verify the effectiveness of the proposed modulation method and circulating current model under the capacitor voltage-balancing control method. In addition, the result also shows that the amplitude of the second circulating current is primarily associated with the arm inductance, and the circulating current will decrease as the arm inductance is increasing.

Keywords MMC-STATCOM · The 2nd circulating current · CPPS (Carrier Pulse Phase Shift)

90.1 Introduction

Nowadays, the static synchronous compensator (STATCOM) has been widely used in the medium-voltage distribution systems for reactive power compensation because of its superior characteristics and competitive costs [1]. At present, the structure based on the modular multilevel converter (MMC) has gradually become a research focus of STATCOM, but there are still many technical difficulties focusing on the establishment of the mathematical model of MMC-STATCOM, the voltage

W. An (✉) · Y. Sun · L. Zhang
Nanjing Institute of technology, 211167 Nanjing, China
e-mail: myfuturewithyou@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_90

regulation of the dc-bus, the dc-link capacitor voltage control, and the suppression of internal circulating current [2–5]; however, the problem of MMC internal circulating current is particularly prominent. The circulating current will increase RMS (Root Mean Square) of the arm current as well as the equipment loss and will even cause transient imbalance and disturbance [6, 7]; in this sense, it is quite essential to calculate the circulating current and suppress it as much as possible.

This chapter proposes a new CPPS modulation strategy and analyzes the internal circulating current of MMC-STATCOM under CPPS; besides, the existence of the second circulating current is proved. Also, the analytical expressions of the second circulating current and the relevant submodule (SM) capacitor voltage ripple are also given in this chapter. Finally, a simulation model is presented to prove the correctness of the derivation.

90.2 MMC-STATCOM Equivalent Circuit

Figure 90.1 shows the equivalent circuit of MMC-STATCOM, where L_0 is the inductance for the bridge arm; R_s and L_s are the ac-side resistance and inductance, respectively; u_{pj} and u_{nj} are the equivalent voltages of the series SMs in upper and lower bridge arms, respectively; i_j is the output AC current; i_d and U_{dc} are the DC current and voltage of MMC-STATCOM, respectively.

90.3 Modulation Strategy

This chapter puts forward a new kind of pulse control mode according to the working principle of MMC, namely the CPPS modulation method. There are two bridge arms in each phase of MMC-STATCOM and the number of the SM in each arm is

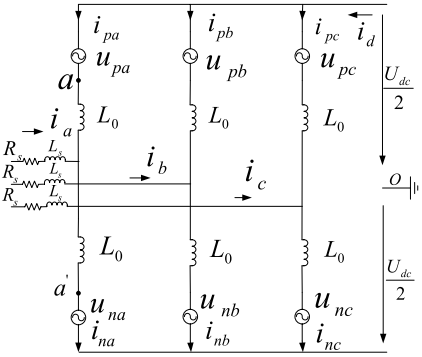


Fig. 90.1 Three-phase MMC-STATCOM equivalent circuit

N. The a-phase is discussed in this chapter because of its three-phase symmetry. The SMs in each bridge arm are numbered from top to bottom, and *i* represents the *i*th SM in the upper arm, as well as *j* represents the *j*th SM in the lower arm. The phase of triangular carriers, corresponding to *N* SMs in the upper bridge arm, is followed by a difference of $2\pi / N$ and they compare with the same sine wave. As a result, the modulation mode in the upper arm is the traditional CPS-SPWM (Carrier Pulse Shift-Sine Pulse Width Modulation), but the pulses of each SM in the lower arm are achieved by the complementary relationship with the SM in the upper arm. If *N* is even and *i* and *j* have the relationship of (90.1), the pulses of *i* and *j* are complementary:

$$2(j-i) = \pm N \ (1 \le i \le N, 1 \le j \le N) \tag{90.1}$$

If *N* is odd and *i* and *j* have the relationship of (2), the pulses of *i* and *j* are complementary:

$$2(j-i)+1 = \pm N \ (1 \le i \le N, 1 \le j \le N) \tag{90.2}$$

The capacitor voltage control method for the MMC-STATCOM under CPPS is shown in Fig. 91.2.

Figure 90.2a is a current decouple control, and Fig. 90.2b includes the voltage-balancing control expressed as V_{Bji} and the average voltage control expressed as V_{Aji} The output variable V_{ji_p} will compare with triangular carriers. A great advantage of this control method under the modulation method of CPPS lies in the fact that only the SM in the upper arm should be controlled and the pulse in the lower arm can be obtained by (90.1) and (90.2); therefore, it will save partial cost in the control system.

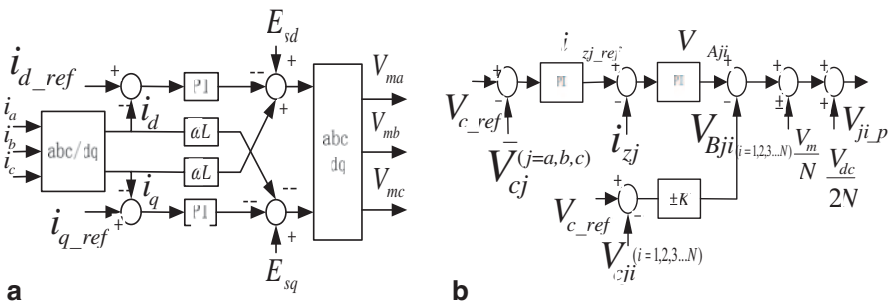


Fig. 90.2 **a** Grid-connected control. **b** Capacitor voltage control method

90.4 Circulating Current Analysis in MMC-STATCOM

There are some assumptions at first:

1. Output current of MMC-STATCOM is sine wave.

$$\begin{cases} U_{sa} = U_M \cos \omega t \\ U_{sb} = U_M \cos(\omega t - \frac{2\pi}{3}) \\ U_{sc} = U_M \cos(\omega t + \frac{2\pi}{3}) \end{cases} \tag{90.3}$$

$$\begin{cases} i_{sa} = I_M \cos(\omega t + \varphi) \\ i_{sb} = I_M \cos(\omega t - \frac{2\pi}{3} \varphi) \\ i_{sc} = I_M \cos(\omega t + \frac{2\pi}{3} + \varphi) \end{cases} \tag{90.4}$$

Where U_M and I_M are the maximum output AC voltage and current, respectively; ω is the fundamental frequency of the AC grid; and φ is the power factor angle.

2. The DC voltage of MMC-STATCOM is constant.
3. The capacitance voltage of SMs in the same bridge arm is equal to each other.

Therefore, the energy distributes averagely in each arm under the above assumptions. The arm's current and voltage are given by the following:

$$\begin{cases} u_{pj} = \frac{1}{2} U_{dc} - U_{sj} \\ u_{nj} = \frac{1}{2} U_{dc} + U_{sj} \\ i_{pj} = \frac{1}{3} i_d - \frac{1}{2} i_j \\ i_{nj} = \frac{1}{3} i_d + \frac{1}{2} i_j \end{cases} \tag{90.5}$$

Where i_d is the dc-side current of MMC-STATCOM; i_{pj} and i_{nj} are the bridge arm currents of the j-phase; u_{pj} and u_{nj} are the bridge arm voltages of the j-phase; $j = a, b, c$.

The a-phase is analyzed, for example, because of the three-phase symmetry. According to (90.3) and (90.4), the output instantaneous power of the a-phase is given by the following:

$$P_{sa}(t) = \frac{1}{2} I_M U_M + \frac{1}{2} I_M U_M \cos 2\omega t \tag{90.6}$$

The instantaneous energy stored in bridge arms is represented as (90.7):

$$\begin{cases} \mathrm{w_{pa}(t)} = \dfrac{1}{2}\mathrm{CNu}\dfrac{2}{\mathrm{cp}}(t) \\ \mathrm{w_{na}(t)} = \dfrac{1}{2}\mathrm{CNu}\dfrac{2}{\mathrm{cn}}(t) \end{cases} \tag{90.7}$$

Where u_{cp} and u_{cn} are the capacitor voltage of SM in the upper and lower bridge arms, respectively, and C is the capacitor of SM. If the number of conduction SMs in the upper bridge arm is m, the conduction SMs in the lower bridge arm will be $N–m$; therefore, u_{cp} and u_{cn} can be given by the following:

$$\begin{cases} \mathrm{u_{cp}} = \dfrac{\mathrm{u_{pa}}}{\mathrm{m}} \\ \mathrm{u_{cn}} = \dfrac{\mathrm{u_{na}}}{\mathrm{N-m}} \end{cases} \tag{90.8}$$

Therefore, the total instantaneous energy stored in the a-phase can be obtained based on (90.5), (90.7), and (90.8). Consequently, the total power in the a-phase can be calculated by the derivation of the instantaneous energy, and it is represented as (90.9):

$$\mathrm{P_a(t)} = -\left(\frac{NC}{2\mathrm{m}^2} + \frac{NC}{2(N-M)^2}\right)\omega \mathrm{U}_M^2 \sin 2\omega \mathrm{t} + \left(\frac{NC}{2\mathrm{m}^2} - \frac{NC}{2(N-M)^2}\right)\omega \mathrm{U_{dc}U_M}\sin \omega \mathrm{t} \tag{90.9}$$

It is obvious that the double-frequency component of (9) and (6) is unequal, which shows there is a double-frequency component in the SM voltage, so is the current. As there is no fluctuation in the dc-bus voltage of MMC-STATCOM, the fluctuation in SMs will be negative with the fluctuation in arm inductances [8]; therefore, the double-frequency fluctuations of SM voltage can be represented as (90.10):

$$U_h \angle \theta + 4\omega \mathrm{LjI}_h = 0 \tag{90.10}$$

where $U_h \cos(\omega t + \theta)$ is defined as $U_h \angle \theta$; θ is the second harmonic voltage phase angle; and I_h is the vector of the second harmonic current.

Therefore, the second circulating current in the arm can be represented as (90.11):

$$I_h(t) = -\frac{U_h}{4\omega \mathrm{L}}\sin(\omega t + \theta) \tag{90.11}$$

Then, the expression of the arm voltage and arm current can be amended as (90.12):

$$\begin{cases} u_p(t) = \dfrac{U_{dc}}{2} - \mathrm{U_M}\ \cos\ \omega\mathrm{t} + \dfrac{m}{N} U_h \cos(2\omega t + \theta) \\ u_n(t) = \dfrac{U_{dc}}{2} + \mathrm{U_M}\ \cos\ \omega\mathrm{t} + \dfrac{N-m}{N} U_h \cos(2\omega t + \theta) \\ i_p(t) = \dfrac{1}{3}\mathrm{i_d} - \dfrac{1}{2}\mathrm{I_M}\ \cos\ (\omega\mathrm{t} + \varphi) - \dfrac{U_h}{4\omega L}\sin(2\omega t + \theta) \\ i_n(t) = \dfrac{1}{3}\mathrm{i_d} + \dfrac{1}{2}\mathrm{I_M}\ \cos\ (\omega\mathrm{t} + \varphi) - \dfrac{U_h}{4\omega L}\sin(2\omega t + \theta) \end{cases} \tag{90.12}$$

So the instantaneous power of the a-phase can be represented as (90.13):

$$P_a(t) = u_p(t) i_p(t) + u_n(t) i_n(t) \tag{90.13}$$

The double-frequency component in P_a(t) can be represented as (90.14):

$$P_{a2}(t) = \frac{1}{3} i_d U_h \cos(2\omega\mathrm{t} + \theta) + \frac{1}{2} U_m I_M \cos(2\omega\mathrm{t} + \theta) - \frac{U_h}{4\omega L} u_{dc} \cos(2\omega\mathrm{t} + \theta) \tag{90.14}$$

Then, the capacitor voltage expression is amended as (90.15):

$$\begin{cases} U_{cp}(t) = \dfrac{U_{dc}}{N} - \dfrac{1}{2\omega c} I_M \sin(\omega\mathrm{t} + \varphi) + \dfrac{U_h}{N}\cos(2\omega\mathrm{t} + \theta) \\ U_{cn}(t) = \dfrac{U_{dc}}{N} - \dfrac{1}{2\omega c} I_M \sin(\omega\mathrm{t} + \varphi) + \dfrac{U_h}{N}\cos(2\omega\mathrm{t} + \theta) \end{cases} \tag{90.15}$$

So the total energy stored in the capacitor is represented as (90.16):

$$W_{ca}(t) = m W_{cp}(t) + (N - m) W_{cn}(t) \tag{90.16}$$

Where $W_{cp}(t)$ and $W_{cn}(t)$ are represented as (90.17):

$$\begin{cases} W_{cp}(t) = \dfrac{1}{2} C U_{cn}^2(t) \\ W_{cn}(t) = \dfrac{1}{2} C U_{cn}^2(t) \end{cases} \tag{90.17}$$

Therefore, the total power in arm capacitors can be obtained by the derivation of $W_{ca(t)}$ and the double-frequency component in it can be represented as (90.18):

$$P_{ca2}(t) = \frac{1}{8\omega C} I_M^2 \sin(2\omega t + 2\varphi) - \frac{2\omega C}{N} U_{dc} U_h \sin(2\omega t + \theta) \tag{90.18}$$

In the same way, the total power in arm inductances can be obtained according to (90.12) and represented as (19):

$$P_L(t) = -\frac{1}{4}\omega L I_M^2 \sin(2\omega + 2\varphi) - \frac{1}{6} i_d U_h \cos(2\omega t + \theta) + \frac{1}{8\omega L} U_h^2 \sin(4\omega t + 2\theta) \tag{90.19}$$

So the double-frequency component in $P_{L(t)}$ can be represented as (90.20):

$$P_{La2}(t) = -\frac{1}{4}\omega L I_M^2 \sin(2\omega + 2\varphi) - \frac{1}{6} i_d U_h \cos(2\omega t + \theta) \tag{90.20}$$

Therefore, the total power in the a-phase can be obtained by adding (90.18) and (90.20), and it can be represented as (90.21):

$$\begin{aligned} P'_{a2}(t) = {} & \left(\frac{1}{8\omega C} - \frac{1}{4}\omega L\right) I_M^2 \sin(2\omega + 2\varphi) - \frac{2\omega C}{N} U_{dc} U_h \sin(2\omega t + \theta) \\ & - \frac{1}{6} i_d U_h \cos(2\omega t + \theta) \end{aligned} \tag{90.21}$$

Because of the equality of (90.14) and (90.21), (90.22), (90.23), and (90.24) can be obtained:

$$\tan = \frac{\begin{aligned} & \left[\left(\frac{1}{8\omega c} - \frac{1}{4}\omega L\right) I_M^2 \cos 2\varphi + \frac{1}{2} U_m I_M \sin\varphi\right] \frac{1}{4} i_d \\ & + \left[\frac{1}{2}\left(\frac{1}{8\omega C} - \frac{1}{4}\omega L\right) I_M^2 \sin 2\varphi - \frac{1}{4} U_m I_M \cos\varphi\right]\left(\frac{1}{4\omega L} - \frac{2\omega C}{N}\right) U_{dc} \end{aligned}}{\begin{aligned} & \frac{1}{2}\left[\left(\frac{1}{8\omega C} - \frac{1}{4}\omega L\right) I_M^2 \cos 2\varphi + \frac{1}{2} U_m I_M \sin\varphi\right]\left(\frac{1}{4\omega L} - \frac{2\omega C}{N}\right) U_{dc} \\ & - \frac{1}{4}\left(\frac{1}{8\omega C} - \frac{1}{4}\omega L\right) i_d I_M^2 \sin 2\varphi + \frac{1}{8} i_d U_M I_M \cos\varphi \end{aligned}} \tag{90.22}$$

$$U_h = \frac{\left[\left(\frac{1}{8\omega c}-\frac{1}{4}\omega L\right)I_M^2\cos 2\varphi+\frac{1}{2}U_m I_M \sin\varphi\right]\left(\frac{2\omega C}{N}-\frac{1}{4\omega L}\right)U_{dc} + \frac{1}{2}\left(\frac{1}{8\omega C}-\frac{1}{4}\omega L\right)i_d I_M^2 \sin 2\varphi - \frac{1}{4}U_m \mathrm{I}_M \cos\varphi}{\cos\theta\left[\left(\frac{1}{4\omega L}-\frac{2\omega C}{N}\right)^2 {U_{dc}}^2 + \frac{1}{4}i_d^2\right]} \tag{90.23}$$

According to (90.11) and (90.23), i_h can be obtained and represented as (90.24):

$$i_h = \frac{\left[\left(\frac{1}{8\omega c}-\frac{1}{4}\omega L\right)I_M^2\cos 2\varphi+\frac{1}{2}U_m I_M \sin\varphi\right]\left(\frac{2\omega C}{N}-\frac{1}{4\omega L}\right)U_{dc} + \frac{1}{2}\left(\frac{1}{8\omega C}-\frac{1}{4}\omega L\right)i_d I_M^2 \sin 2\varphi - \frac{1}{4}i_d U_M \mathrm{I}_M \cos\varphi}{4\omega L\cos\theta\left[\left(\frac{1}{4\omega L}-\frac{2\omega C}{N}\right)^2 {U_{dc}}^2 + \frac{1}{4}i_d^2\right]} \tag{90.24}$$

It is obvious that based on (90.24), i_h is the inherent current of MMC-STATCOM and cannot be eliminated by the control method, which means i_h is the minimum of the arm circulating current.

90.5 Simulation Results

In order to verify the correctness of (90.22), (90.23), and (90.24), MMC-STATCOM model applied to a 10-kV system is built, and the number of SMs in each phase is 16. The system parameters are as follows: $R_S = 0.6\ \Omega$, $L_S = 4$ mH, $C = 6$ mF, $Q = 2$ MVar.

Figures 90.3, 90.4, and 90.5 are the results of L_o=32 mH. Figure 90.3 shows that the phase voltages U_S and U_{pcc} are in phase. Besides, Fig. 90.4 shows that the arm current is close to sine wave because i_h is quite small. Figure 90.5 is the second circulating current which is near to the calculated results; therefore, the simulation results can prove the correctness of the above derivation. Figure 90.6 displays that the second circulating current will decrease when the arm inductance is increasing.

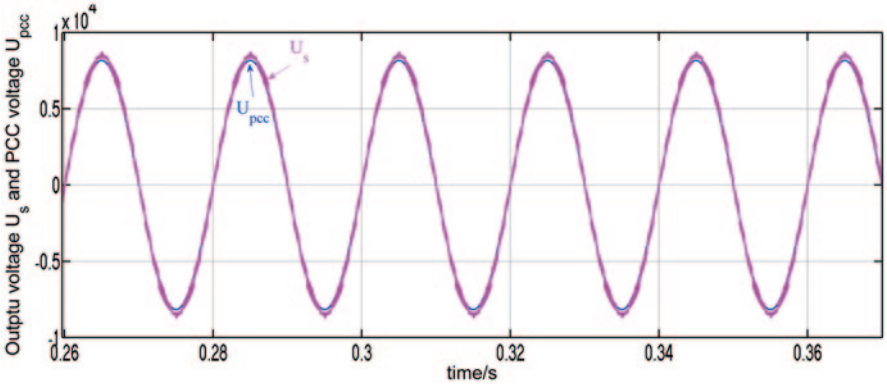


Fig. 90.3 Output voltage U_s and PCC (Point of Common Coupling) voltage U_{pcc}

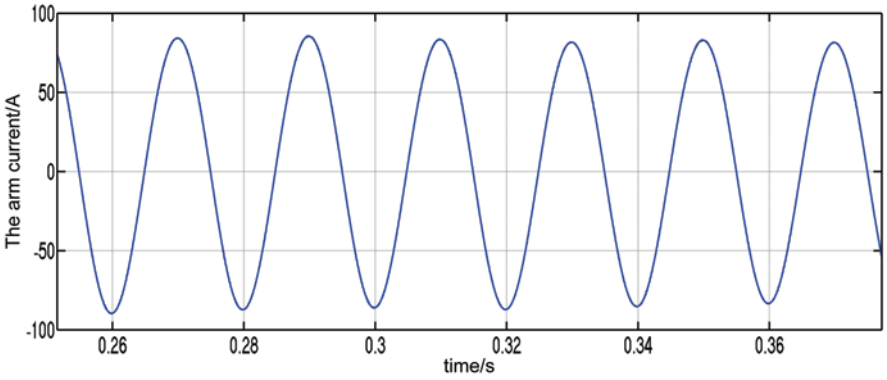


Fig. 90.4 The arm current

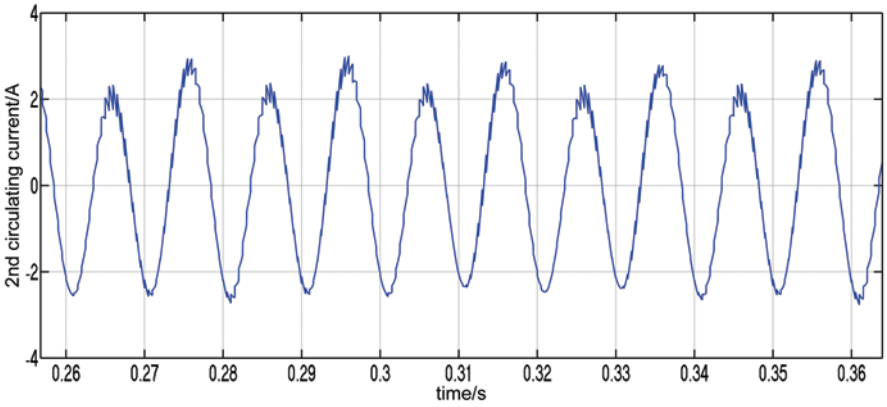


Fig. 90.5 The second circulating current i_h

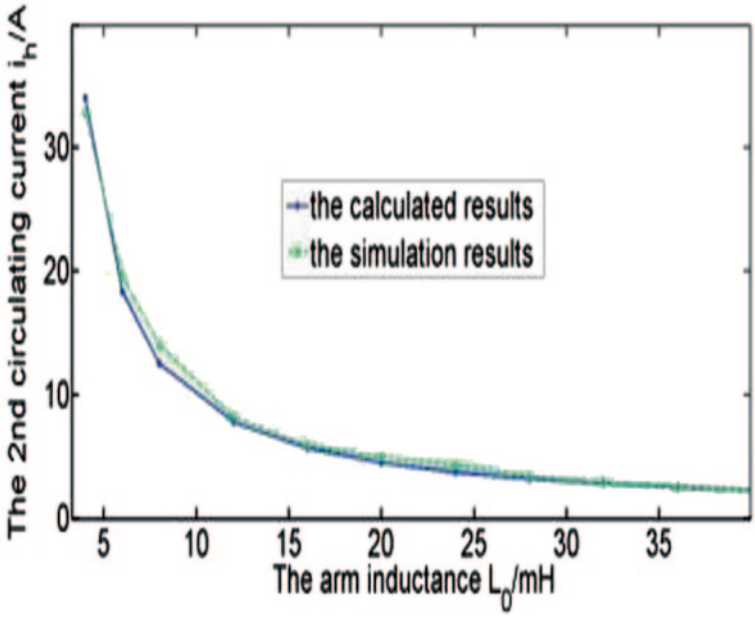


Fig. 90.6 i_h under different inductances

90.6 Conclusion

In this chapter, a new MMC-STATCOM modulation method, CPPS, is proposed, and the relevant control method is introduced. It can guarantee that the conduction number of SM is N at any time, while it can also simplify the control system at the same time;besides, the existence of the second circulating current is proved, and the analytical expressions of the second circulating current as well as the relevant SM capacitor voltage ripple are derived. The simulation results have proved the correctness of the derivation and its impact on the circulating current of arm inductance.

References

Pereira M, Retzmann D, Lottes J, Wiesinger M, Wong G. SVC PLUS: An MMC STATCOM for Network and Grid Access Applications. IEEE Trondheim Power Tech Conf, pp. 1–5; 2011.

Chao M, Wang Y, Liu P, Chong W. Model predictive control of modular multilevel converter-based STATCOMs. Elect Drive. 2013;S1:52–6. (in Chinese).

Dai K, Chen X, Ding Y, Kang Y. Research on modeling and hierarchical control for an improved MMC-STATCOM. Trans China Electrotechnical Soci. 2013;28(12):44–51. (in Chinese).

Lian X, Jiang D. Fault-tolerant design and control of modular multilevel converter-based STATCOM. Power Syst Protection Control. 2012;40(16):83–9. (in Chinese).

Liang H. Research of the control strategy on modular multilevel converter based STATCOM. Hefei: Hefei University of Technology; 2013. (in Chinese).

Chungang C, Zhao C, Chen X. Mathematical model and control strategy of MMC-HVDC. Proc CSU-EPSA. 2012;24(4):13–8. (in Chinese).
Yang V. Research on modular multilevel converter (MMC). Beijing: Beijing Jiaotong University; 2011. (in Chinese).
Li Z, Wang P, Chu Z, Zhu H, Luo Y, Li Y. An inner current suppressing method for modular multilevel converters. IEEE Trans Power Electron. 2013;28(11):4873–9.

Chapter 91
A Line Structured Light Visual Sensor for Road Obstacle Detection

Haiyan Shao, Zhenhai Zhang, Kejie Li, Jian Wang, Tao Xu, Shuai Hou and Liang Zhang

Abstract A new line structured light visual sensor is proposed. The line structured light for remote measurement is generated by high power semiconductor laser point source with miniature piezo fast steering mirror. The mathematical model of the line structured light sensor and calibration method is described. By analyzing the line profiles of structured light under different obstacles of road, the mathematical model of line structured light is obtained. Finally, the possibility of using the line structured light visual sensor to detect road obstacle is proposed.

Keywords Line structured light visual sensor · Piezo steering mirror · Line profile of structured light · Obstacle detection

91.1 Introduction

Research on structured light vision measuring method originated in the 70's of the last century. It's developed from optical triangulation method, and is now widely used in the field of industry [1]. Line structured light visual sensor consists of one laser projector and one camera. The laser produces some kind of structured light by optical lens, and projects to the measured objects. And the light stripe is modulated by the depth of the measured objects. The camera captures the image of the measured object with the deformed structured light stripe. Then the 3D characteristic information of the measured object surface can be acquired by optical triangulation method from the 2D deformed light stripe image. Road obstacle detection by

Z. Zhang (✉) · H. Shao · K. Li · J. Wang · T. Xu · S. Hou · L. Zhang
School of Mechatronic Engineering, Beijing Institute of Technology, 100081 Beijing, China
e-mail: zhzhang@bit.edu.cn

H. Shao
School of Mechanical engineering, University of Jinan, 250022 Jinan, China

H. Shao
e-mail: me_shaohy@ujn.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_91

structured light is still on the stage of theoretical study and experimental design, and is not mature in application at home [2–3], which is primarily because of unpredictable obstacle variety, features recognize difficult and remote security detection means lack.

In this paper, a line structured light vision sensor used for remote detection is established. The mathematics model of structured light strip is analyzed.

91.2 Mathematical Model of Line Structured Light Vision Sensor

The line structured light is projected from the laser point source by piezo steering mirror. Laser feature the advantages of high brightness and good direction etc. than ordinary light. The laser point light source can greatly reduce the power than the line source with more security. The schematic of line structured light vision sensor based on laser point light is shown in Fig. 91.1.

Regard the camera coordinate $(O_cX_cY_cZ_c)$ as the vision sensor coordinate. Assume the transformation coordinate is $(O_rX_rY_rZ_r)$. Then the relationship of transformation coordinate and camera's can be described by rotation matrix R and translation matrix T (Equation 91.1). Matrix R and T can be calculated by the sensor calibration.

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = R \begin{pmatrix} X_r \\ Y_r \\ Z_r \end{pmatrix} + T \tag{91.1}$$

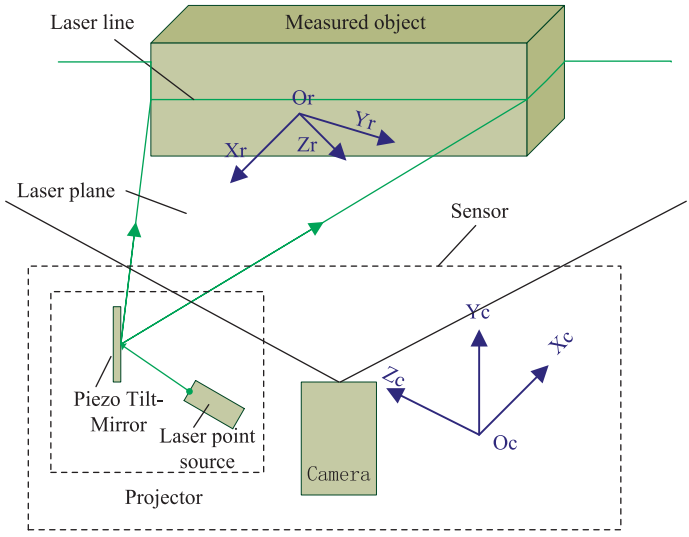


Fig. 91.1 Schematic of line structured light vision sensor

where $R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$, $T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$.

The mathematic model of structured light vision sensor is derived from the imaging model of two-dimensional camera [4].

$$\begin{cases} X_c = \dfrac{(x - x_0 + \Delta x)}{c} Z_c \\ Y_c = \dfrac{(y - y_0 + \Delta y)}{c} Z_c \\ Z_c = f(X_c, Y_c) \end{cases} \tag{91.2}$$

where: c is the effective focus of camera; (x, y) is the imaging-point coordinate of measured point in the camera; $(x_0, y_0)$is the image plane center; $(\Delta x, \Delta y)$ is the integrated imaging distortion.

If we select one point in light plane as the reference coordination origin O_r, make coordinate $(X_r\ Y_r)$ and light plane coincidence, and assume Z_r axis meet the right-handed coordinate system. Then the equation of light plane in reference coordination is

$$Z_r = 0 \tag{91.3}$$

According to Equation (91.1), the mathematical model of the line structured light sensor is substituted in Equation (91.2):

$$\begin{cases} X_c = \dfrac{(x - x_0 + \Delta x)}{c} Z_c \\ Y_c = \dfrac{(y - y_0 + \Delta y)}{c} Z_c \\ Z_c = \dfrac{r_{22}r_{31} - r_{21}r_{32}}{r_{11}r_{22} - r_{12}r_{21}} X_c + \dfrac{r_{11}r_{32} - r_{12}r_{31}}{r_{11}r_{22} - r_{12}r_{21}} Y_c + \dfrac{r_{21}r_{32} - r_{22}r_{31}}{r_{11}r_{22} - r_{12}r_{21}} t_1 \\ \quad + \dfrac{r_{12}r_{31} - r_{11}r_{32}}{r_{11}r_{22} - r_{12}r_{21}} t_2 + t_3 \end{cases} \tag{91.4}$$

For a multi-line structured light vision sensor, it can be seen as the extension of single line sensor. The mathematical model is similar to the single line one. The difference is laser projector launched i light planes (where i is the number of structured lights).

91.3 Image Processing

91.3.1 Models for Line Structured Light Profiles

The profile of straight line in 2D image appears line in 1D. The ideal light strip general has certain width, and is assumed to have a symmetry strip profile in the normal direction [5]; however, the sampling effects of the lines often don't have this profile. Figure 91.2 shows a typical profile of a structured light in an image. The gray change of transition region on the two sides is not a step change, but transits slowly, similar to parabolic curve; thus it's more suitable for the parabolic model to describe the profile in Equation (91.5).

$$f_p(x) = \begin{cases} h(1-(x/w)^2) & |x| \le w \\ 0 & |x| > w \end{cases} \tag{91.5}$$

where: x is the off center distance in the normal direction; w is the light-strip width; h is the gray value of light-strip; $f_p(x)$ is the simulation curve.

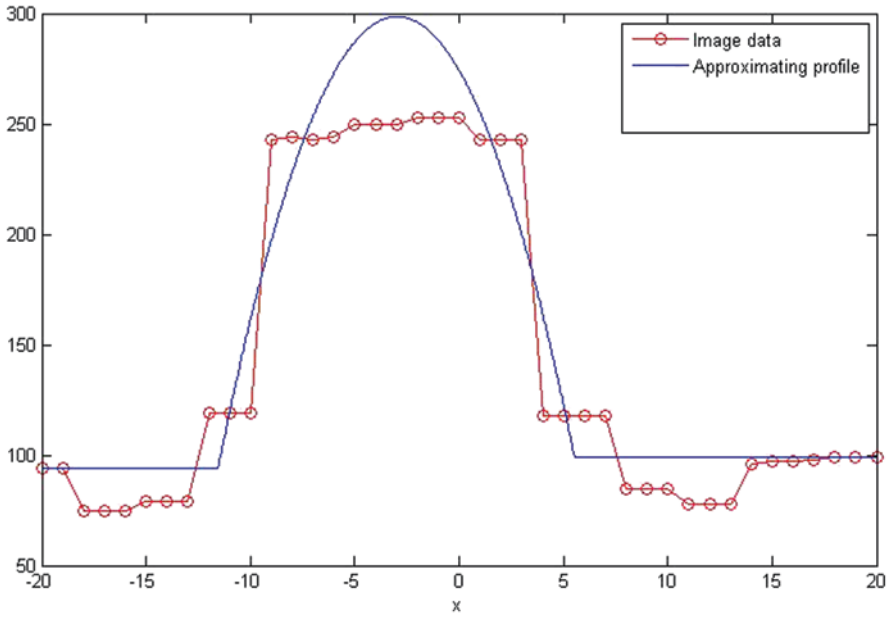


Fig. 91.2 Profile of a line in a structured light image and approximating parabolic line profile

91.3.2 Image Processing

Firstly, use Roberts operator for the edge detection. The method uses the partial differential operators for edge. The edge location precision is higher. The operator for a steep edge image effect and low noise is better. Figure 91.3b shows the edge of the concrete pier clearly. Because of the image without smoothing, it does not have the capability of anti noise.

Then, use the mean filter of neighborhood average method to remove particulate noise in images (Fig. 91.3c). The core idea of neighborhood average method is mean gray processing [6]. For each of the pixels (m,n) in given image $f(i,j)$, select the neighborhood S. Suppose S contains of M pixels, taking the mean value as the gray value of (m,n). In a word, use the grayscale average of neighborhood pixel gray values instead of the pixel's original value. Result is shown in Fig. 91.3.

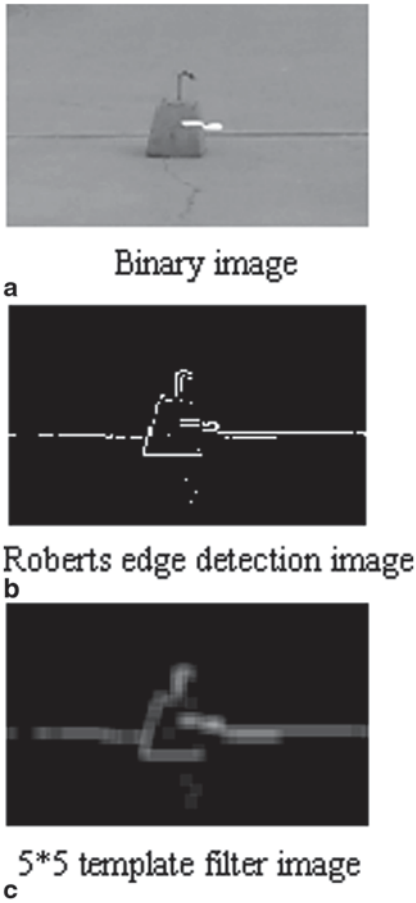


Fig. 91.3 Image processing of light strip **a** Binary image. **b** Robert operator edge detection image. **c** 5*5 template filter image

Fig. 91.4 Experimental equipment

91.4 Experiments

We design a line structured light sensor, as showed in Fig. 91.4. The sensor is consisted of a camera, a laser point source and a piezo steering mirror. The camera is multi-spectral prism AD130GE. The camera lens is equipped a 532 nm ± 5 nm narrow band filter to suppress background light. Laser projector consists of a laser point source and a piezo steering mirror. Laser projector is continuous point source of pumped solid state laser, wavelength 532 nm, power 1.8 w, initial spot diameter 8 mm and beam divergence 0.4 mrad. The piezo steering system consisted of piezo steering mirror and double channel control box. Piezo steering mirror used PI S334 [7–8]. It's a miniature fast steering mirror with up to 120 mrad (~6.8°) deflection, mirrors diameter 10 mm, millisecond response and resolution to 0.2 μrad, R > 98% and λform 500 nm to 2 μm. Sigma Koki XYZ 3 axis adjusting platform is arranged below the point laser projector. Manfrotto gear head is arranged below PI steering mirror, after pitching + 90/− 30° and around the pitch + 30/− 90°.

Project the structured light respectively to concrete pier, carton and brick wall. The section profile of the light strip is shown in Fig. 91.5. The result is consistent with the previous analysis. Figure 91.5 shows that the surface of concrete pier and carton is smooth, so the structured light is distributed with the boundary more obvious than brick wall.

91.5 Conclusions and Future Work

The line structured light for remote measurement is built by high power semiconductor laser point source with miniature piezo fast steering mirror. The experiment shows the line structured light vision sensor could detect and recognize the interesting concave convex obstacle far car 50 m.

The next step will be to realize remote detection of road obstacle so as to fuse line structured light sensor identification information and LIDAR information, achieve more accurate and integrated detection of road obstacles detection at all time and all weather.

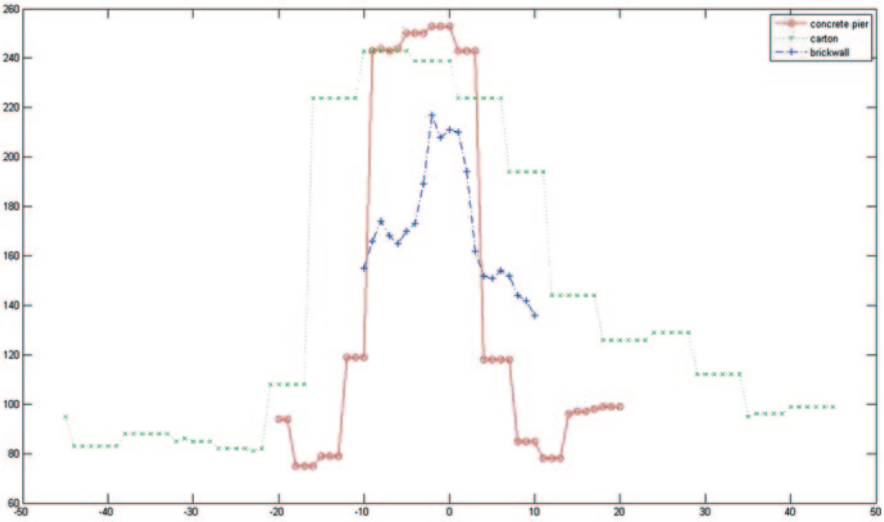


Fig. 91.5 Line structured light strip section profile of different obstacles

Acknowledgements The work is supported by the National Natural Science Foundation of China (61273346), the National Defense Major Fundamental Research Program of China (C××2011××03), the National Defense Key Fundamental Research Program of China (A××2011××05 & A××2013××10), the Specialized Research Fund for the Doctoral Program of Higher Education (20121101120009), the Scientific Research Foundation for the Returned Overseas Chinese Scholars, Excellent Young Scholars Research Fund of Beijing Institute of Technology (2012YG0203), and the Program for the Fundamental Research of Beijing Institute of Technology (2013CX04031, 2011CX02023 & 2014CX02031).

References

1. Will PM, Pennington KS. Grid coding: a preprocessing technique for robot and machine vision [J]. Sess No. 3 Scene Anal II Gen Pap. 1971;2(3):319.
2. Li L, Sun L. Line-structured light image processing procedure for pavement rut detection [J]. J Tongji univ (Nature Science). 2013;41(5):710–5 (In Chinese).
3. Nanjing University of Science and Technology. Intelligent 3D laser road state detecting vehicle. China: ZL 03113468.8[P]; 2004. p. 11–24(In Chinese)
4. Hao J, Yu Z. Vision measurement principle and method [M]. Beijing: China Machine Press; 2011 p. 6–168.(In Chinese)
5. Carsten S. An unbiased detector of curvilinear structures [J]. IEEE Trans Pattern Anal Mach Intell. 1998;20(2):113–25
6. Bai F. Vision measurement technology [M]. Beijing: Publishing House of Electronics Industry; 2013. p. 8–32(In Chinese)
7. Xu F, Ji M, Zhao C. Status of fast steering mirror [J]. J Appl Opt. 2010;31(5):847–50. (In Chinese)
8. Physik Instrumente. Piezo Tip-Tilt Platforms. http://www.physikinstrumente.com/en/products/steering_mirror/index.php Accessed: 21 June 2014.

Chapter 92
A Method to Eliminate the Effect of Sound Path Length Change of Ultrasonic Heat Meter

Kai Mao, Yan Chai, Jiejie Shao, Liang Hu, Peng Ye and Xin Fu

Abstract Ultrasonic heat meter (UHM) plays a key role in the heat consumption sensing in automatic metering and control system of district heat supply pipe network; however, the large temperature changes (below 10℃ to above 90℃) of water in the network generally cause sound path length (SPL) changes due to thermal expansion and contraction, which will greatly reduce the accuracy of UHM in practice. This paper proposes a method to eliminate the effect of SPL change for UHM. In the method, a novel model has been developed to calculate flow rate and then heat consumption from upstream and downstream transmit times and sound velocity, but not containing the SPL parameter commonly existing in traditional models. A sample learning method is proposed in addition to correct the effect of sound velocity changes under different temperatures. The method is validated on actual UHMs by flow rig experiments on simulating serious temperature variations.

Keywords Ultrasonic heat meter · Transit time difference · Sound path length · Temperature change

92.1 Introduction

The automatic metering and control systems composed by ultrasonic heat meters (UHMs), control valves and remote data transmission modules are widely used in district heat supply pipe networks. In the systems, the UHMs play key roles for heat consumption sensing. They generally measure the flow rates of hot water supplied into the users' rooms based on differences between the upstream and the downstream transit times of the ultrasonic pulses [1], and then calculate the heat

L. Hu (✉) · K. Mao · Y. Chai · J. Shao · X. Fu
State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, 310058 Zhejiang, China
e-mail: cmeehuli@zju.edu.cn

P. Ye
Tancy Group Co. Ltd, 310058 Zhejiang, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_92

consumption which combines the flow rates with water the temperatures incoming and outgoing the rooms.

The transit time of the ultrasonic pulses in UHM is determined by factors including the sound path length (SPL) between the transmitter and receiver transducers, the sound velocity and the flow velocity of the water. The change of SPL will directly affect the measurement accuracy of UHM [1–4]. This kind of change may be caused by the pipe scaling due to the dirty of the pipe water [3]. Another reason lies in the thermal expansion and contraction of the measuring pipe on which the transducers are installed because the water temperature in the pipe network may changes from below 10ºC to above 90ºC in actual use [4]. In addition, the water temperature changes will also affect the sound velocity and causes additional errors of the flow rate determination.

In order to reduce the effects of sound velocity variation and the SPL variation, the traditional method eliminates the sound velocity parameter in the calculating formula [5, 6] and calibrates the UHMs at different temperatures to obtain according instrument coefficients so as to achieve the compensation of the SPL changing effects; however, the SPL changing characteristics of each UHM is different due to the manufacture and assembling differences of the measurement pipe. Thus it is different to compensate the effects of SPL changes in practice. In some industrial ultrasonic flow meters, manufacturers measure the SPL of each measuring pipe in previous and take advantage of well characterized material temperature properties of the pipe to achieve temperature compensation [6]. This method will take a lot of time to produce a qualified flow meter; as a result, it is not applicable for mass production of UHMs. Artificial Neural Network was also proposed in industrial ultrasonic flow meters to examine the temperature and pipe diameter variation [5]. This method put forward a very high request to computing power of flow meter. In order to acquire lower power consumption, it is also not applicable for UHMs to adopt this method.

This paper proposes a new method to eliminate the effect of SPL change for UHM. In the method, a novel model has been developed to calculate flow rate and then heat consumption from upstream and downstream transmit times and sound velocity, but not containing the SPL parameter commonly existing in traditional models. A sample learning method is proposed in addition to correct the effect of sound velocity changes under different temperatures. The method was validated on actual UHMs by flow rig experiments simulating serious temperature variations.

As the most important advantage of this method, the accomplishment of the correcting works of UHMs with the same DN just needs a few UHMs for sample leaning. What's more, this method can be adopted in any kind UHMs without any hardware modification and any power consumption addition.

This paper is organised as below: introduction in Sect. 1, which is followed by a detail description of the compensation method of the changes of the SPL and the sound velocity in the UHM. Sect. 3 verifies the proposed method in this paper by using an experiment. Finally, the conclusion is given in Sect. 4.

92.2 Method

92.2.1 Measurement Principle

The UHM generally combines a pair of ultrasonic transducers installed upstream Tu and downstream Td of the measurement pipe with a certain angle θ as shown in Fig. 92.1. The UHM measures flow rate with the transit time difference Δt of the sound velocity C between the upstream and downstream within a certain SPL l.

The transit-times of upstream $t_{u\text{-}d}$ and downstream $t_{d\text{-}u}$ can be calculated by the Eq. 92.1 and Eq. 92.2 respectively [7, 8].

$$t_{u-d} = \frac{l}{C + u\cos\theta} \tag{92.1}$$

$$t_{d-u} = \frac{l}{C - u\cos\theta} \tag{92.2}$$

In order to reduce the effect of the sound velocity C, the traditional method uses the Eq. 92.3. Obviously, there is a no relationship between the sound velocity C and the media velocity u in the Eq. 92.3 [5, 6], but is still affected by SPL l.

$$u = \frac{(t_{d-u} - t_{u-d}) \bullet l}{2(t_{d-u} \bullet t_{u-d}) \bullet \cos\theta} = k\frac{t_{d-u} - t_{u-d}}{t_{d-u} \bullet t_{u-d}} \tag{92.3}$$

The k is the instrument coefficient; however, the SPL changing characteristics of each UHM is different due to the manufacture and assembling differences of the measurement pipe; thus it is different to compensate the effects of SPL changes in practice.

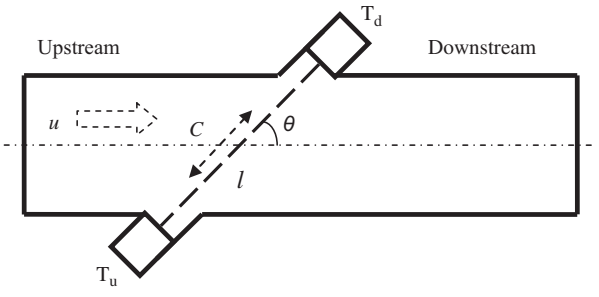


Fig. 92.1 Measuring principle of the transit time ultrasonic heat meter

92.2.2 Calculation Model to Eliminate the Effect of Sound Path Length Changes

Taking the difficulty of compensating the effects of the SPL changes into account, this paper proposes a method to eliminate the effect of SPL change for UHM. In the method, a novel calculation model which doesn't contain the SPL parameter commonly existing in traditional models was developed. The related equations are shown in Eq. 92.4~Eq. 92.7.

$$\Sigma t = t_{d-u} + t_{u-d} = \frac{2lC}{C^2 - u^2 \cos^2 \theta} \tag{92.4}$$

$$\Sigma t = t_{d-u} + t_{u-d} = \frac{2lC}{C^2 - u^2 \cos^2 \theta} \tag{92.5}$$

$$\frac{\Delta t}{\Sigma t} = \frac{2lu \cos\theta}{C^2 - u^2 \cos^2 \theta} \frac{C^2 - u^2 \cos^2 \theta}{2lC} = \frac{u \cos\theta}{C} \tag{92.6}$$

$$u = \frac{\Delta t \cdot C}{\Sigma t \cdot \cos\theta} \tag{92.7}$$

Eq. 92.7 is the calculation model which doesn't contain the SPL parameter, obviously, the effect of SPL changes is eliminated in theory; thus in order to calculate the media velocity u, just two parameters, the transit-times of upstream $t_{u\text{-}d}$ and downstream $t_{d\text{-}u}$, should be measured. It is easy to adopt a single chip microcomputer and a clock chip whih are both low power consumption to get the above two parameters. With this method, the request of low power consumption for UHM is also achieved at the same time.

In practice, the manufacture precision of pipes and the installation precision of transducers are very low, which leads the consistency is very bad between different UHMs; as a result factories of UHM adopt the real flow calibration to calibrate the measurement accuracy commonly. Then Eq. 92.7 is changed to be Eq. 92.8 simply.

$$u = \frac{\Delta t \bullet C}{\Sigma t \bullet \cos\theta} = k' \frac{\Delta t}{\Sigma t} \tag{92.8}$$

The $k' = C/\cos\theta$ is named as the instrument coefficient. In order to get an accurate, an original value is usually set before calibration, and then compares the measured value of water volume by UHM with the standard value of water volume got by real flow calibration device to acquire the actual use value. Thus the instrument coefficient k' reflects the measurement accuracy of UHM.

92.2.3 Sample Leaning to Correct the Effect of Sound Velocity Changes

Obviously, the instrument coefficient *k'* is affected by the sound velocity *C*. The rules of the sound velocity *C* vary with the temperature *T* changes, which can be described as Eq. 92.9 [9].

$$C = 1404.3 + 4.7\ T - 0.04\ T^2 \tag{92.9}$$

According to Eq. 92.9, the rule is approximate linearity. In order to ensure a high production efficiency and measurement accuracy, we adopt a sample leaning method which corrects the effect of sound velocity changes segmentally to acquire different correction factors at different temperature ranges, for example at 0℃ ~ 10℃, 10℃ ~ 20℃, etc. Then use the same factors to calculate the other instrument coefficients of UHMs which are the same batch and have the same DN. The segmental correction formula of instrument coefficient is shown in Eq. 92.10.

$$k'' = k'\left[1 + (T_2 - T_1) \bullet \xi\right] \tag{92.10}$$

T1 is the basic temperature and *k'* is the basic instrument coefficient corresponding to *T1*.

T2 is the target temperature and *k''* is the target instrument coefficient corresponding to *T2*.

ξ is the correction factor.

92.3 Experiment

The experiment is taken on a heat meter real flow calibration table which adopted the weight method. The experiment table commonly contains the following elements: big water tank, small water pump, PT1000, electronic balance, valve and pipe, as shown in Fig. 92.2.

The work principle is simple. Heaters are used to heat the water in the big water tank that can simulate the practical temperature condition. The water pump is used to transport the water in the big water tank to flow through UHMs into the small water tank which has an electronic balance under the button to weigh the water quality. Then combine the above quality of water and the corresponding density which according to the temperature measured by PT1000 so as to calculate the water volume which flows through the heat meter. The corresponding actual experiment table is shown in Fig. 92.3.

The first step of the experiment is to get the value of *k'* in Eq. 92.10 at the corresponding temperature *T1*. In most practical applications of district heat supply pipe network, the water temperature is near 55℃, so at this experiment we choose

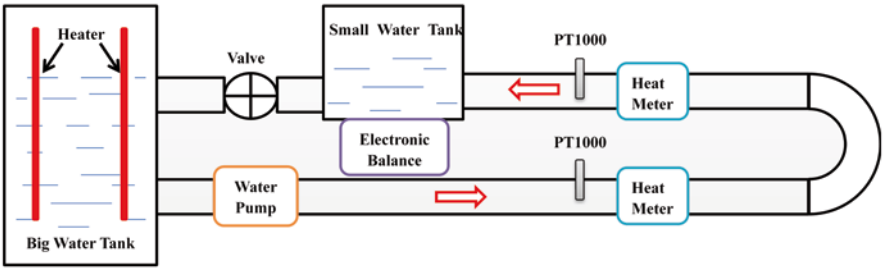


Fig. 92.2 Work principle block diagram of heat meter calibration table

Fig. 92.3 Actual experiment table. **a** real flow calibration table. **b** ultrasonic heat meters

T_1=55ºC. It is easy to calculate the value of *k'* at the T_1=55ºC by comparing the measured value by UHM with the standard value gotten by the real flow calibration. In addition, *k'* is an array actually that contains three arrays (k_1', k_2' and k_3') corresponding to three different flow rates as defined in the UHM standard, because the difference of instrument coefficients is big and nonlinear at different flow rates. So k" is also an array containing three arrays (k_1", k_2" and k_3").

The second step is to get accurate instrument *k"* at corresponding temperature range by real flow calibration. The data of one meter at different temperature ranges which we acquire at this experiment are shown in Table 92.1, including the *k"* and the corresponding accuracy.

The third step is to use Eq. 92.10 to calculate the corresponding correction factors ξ which also is an array comprised of ξ_1, ξ_2 and $\xi_{3\ at}$ different flow rates. Of course, in order to let the ξ fit most UHMs, more *k"* tables of different meters are needed. In this paper, we adopt eight different meters to calculate appropriate ξ values in Table 92.2.

In the end, we select a UHM randomly to compare the k_m" measured by the ultrasonic heat meter calibrating device as the accurate values with the k_c" calculated by the method proposed in this paper. The dotted lines represent the k_m" and the solid linesrepresent the k_c" as shown in Fig. 92.4.

Table 92.1 The k'' values of one meter at different temperature ranges and flow rates with the corresponding measurement accuracy

Media temperature	Big flow (5.0 m^3/h)		Middle flow (2.5 m^3/h)		Small flow (0.5 m^3/h)	
	k_1''	Accuracy (%)	k_2''	Accuracy (%)	k_3''	Accuracy (%)
20ºC	1717	0.16	1781	0.07	1984	−0.12
25ºC	1693	0.08	1754	−0.43	1930	0.28
35ºC	1677	−0.22	1738	−0.23	1898	0.18
45ºC	1655	0.05	1700	−0.06	1860	−0.63
55ºC(k')	1636	0.00	1679	−0.08	1805	−0.16
65ºC	1618	−0.04	1658	−0.30	1783	0.43
75ºC	1609	0.05	1648	0.21	1788	−0.40

Table 92.2 ξ Values at different temperature ranges and flow rates

Temperature range	Big flow (5.0 m^3/h)	Middle flow (2.5 m^3/h)	Small flow (0.5 m^3/h)
	ξ_1	ξ_2	ξ_3
20–25ºC	0.0015	0.0016	0.0027
25–35ºC	0.0013	0.0013	0.0026
35–45ºC	0.0011	0.0012	0.0025
45–55ºC	0.0010	0.0012	0.0015
55–65ºC	0.0000	0.0000	0.0000
65–75ºC	0.0010	0.0012	0.0012
75–80ºC	0.0009	0.0010	0.0005

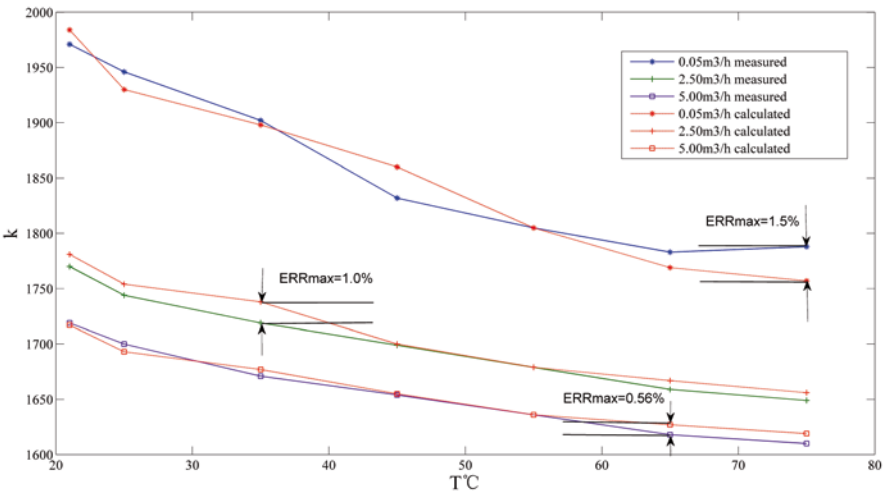


Fig. 92.4 Errors between k_m'' and k_c'' at different temperatures and rates

The errors of the UHM can be represented by the error between the k_m'' and k_c''. In accordance with the standard of UHM [10], the errors should be less than 2.0 % at the 2.5 m^3/h and 5 m^3/h flow rates and the error just need to be less than 3.0 % at the 0.5 m^3/h flow rate. As shown in Fig. 92.4, the maximum error are 1.5 % 1.0 % and 0.56 % at above three flow rates respectively. So using this method at this experiment, the measurement errors conform to the standard totally.

In addition, the production efficiency is improved significantly. With the traditional method in Eq. 92.3, the method of acquiring the instrument coefficient k for each UHM is using real flow calibration at different temperatures and flow rates. Ignoring the calibration time, it takes near 24 h to heat the water in water tank from the room temperature to 90ºC then cool it to the room temperature naturely. Compared with the traditional method, the method proposed in this paper in the same batch and DN of UHMs will not need to adjust the water temperature. As a result, the production efficiency will be improved significantly.

92.4 Conclusion

This paper proposes a method to eliminate the effect of SPL change for UHM. In the method, a novel model was developed to calculate the flow rate from upstream and downstream transmit times and sound velocity, but not containing the SPL parameter commonly existing in traditional models. A sample learning method is also proposed in addition to correct the effect of sound velocity changes under different temperatures. The method has been validated on actual UHMs by flow rig experiments simulating serious temperature variations. With this method, not only the accuracy conforms to the standard of UHM, but also the production efficiency has been improved significantly while compared with the traditional.

Acknowledgements The authors are grateful to the financial supports of the National High Technology Research and Development Program (863 program) of China (No. 2012AA040605), the National Natural Science Foundation of China (No. 51105329) and the Special Support from Postdoctoral Science Foundation of China (No. 2012T50509).

References

1. Baker RC. Flow measurement handbook. Cambridge University: Cambridge University Press; 2005. pp. 330.
2. Drenthen JG, Kurth M, Hollander HD, Vermeulen M. Reducing installation effects on ultrasonic flow meters. Krohne CT Products & Krohne Oil & Gas, 27th North Sea Flow Measurement Conference; 2009. pp. 1–18.
3. Hogendoorn J, Hofstede H, Brakel PV, Boer A. How Accurate are Ultrasonic Flowmeters in Practical Conditions Beyond the Calibration. 29th International North Sea Flow Measurement Workshop, p, 25–8 October 2011, Tonsberg, Norway; 2011.
4. Tawackolian K, Buker Or, Hogendoorn J, Lederer T. Calibration of an ultrasonic flow meter for hot water. Flow Meas Instrum. 2013;30(1):166–73.

5. Santhosh KV, Roy BK. An intelligent flow measurement technique using ultrasonic flow meter with optimized neural network. Intern J Control Autom. 2012;5(4):185–95.
6. McDonald BE, Sui L. Ultrasonic flow measurement with integrated temperature measurement compensation. 16th international flow measurement conference 2013.ge oil & gas-measurement & control, IMEKO-International Measurement Federation Secretariat. 2013 pp 570–5.
7. ANSI/ASME MFC-5M-1985. measurement of liquid flow in closed conduits using transit-time ultrasonic flowmeters, an american national standard; 1994.
8. Lynnworth LC. Ultrasonic measurement for process control: theory, techniques, aplications. London: Academic Press; 1989. pp 10–5.
9. Lubbers J, Graaff R. A simple and accurate formula for the sound velocity in water. Ultrasound Med Biol. 1998;24(7):1065–8.
10. prEN 1434–1 2006. Heart Meters. CEN/TC 176; 2006.

Chapter 93
A Novel Ultrasonic Transducer Module with an Own Reference Sound Channel for Flow Measurement

Jun He, Weilong Guo, Jiejie Shao, Kai Mao, Liang Hu, Peng Ye and Xin Fu

Abstract Zero error caused by transducer system asymmetry which drifts with the environment change is a severe limitation to the performance improvement of an ultrasonic flow meter. This chapter proposes a novel ultrasonic transducer module with an own reference sound channel for zero error calibration. Different from traditional transducers, the module consists of two energy conversion terminals which are connected by a reference sound channel. Part of the ultrasound generated by one terminal transfers into the flow media and finally into another terminal, which is used for flow measurements. The rest propagates through their own sound channel as reference signal for zero drift acquisition. According to the difference of the measuring signal and the reference signal, the zero drift error can be calibrated. We have illustrated the construction, zero elimination method, and design theory of the module with simulation employed for validation. A prototype has been developed and the preliminary results demonstrate an observable decrease of the zero drift error.

Keywords Transducer module · Reference sound channel · Zero drift error · Ultrasonic flow meter

93.1 Introduction

The ultrasonic flow meter (UFM) has been widely used because of its advantages including no-moving-parts construction, no pressure loss, and bidirectional measurement. A large number of UFMs belong to the transit time type whose principle is based on the difference between the transit times of ultrasonic pulses emitted over one or more propagation paths upstream and downstream of the moving fluid [1]. It has been demonstrated that the performance of this kind of UFMs is quite

L. Hu (✉) · J. He · W. Guo · J. Shao · K. Mao · X. Fu
State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, 310058 Zhejiang, China
e-mail: cmeehuli@zju.edu.cn

P. Ye
Tancy Group Co. Ltd, 310058 Zhejiang, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_93

interrelated with the reciprocity of the system [2, 3]. Under reciprocal operation, i.e., the transducer systems are absolutely symmetrical, the signal waveforms in both directions remain consistent and the time of flight only depends on the flow status. Otherwise, the difference arises and will cause the zero drift error which is a severe limitation to the performance improvement of UFMs.

A study has shown that the error can be eliminated by achieving either "perfect transducer symmetry" (both the transducers used are identical) or "perfect electrical symmetry" (the impedance of the transmitting circuit equals to the one of the receiving circuit) [4, 5]. However, it is apparently impractical to ensure the absolute consistency in transducer productions [6]. Due to the demand difference of transmitting and receiving circuits, the perfect electrical symmetry is also hard to realize. Thus, a commonly used method is to calibrate the zero drift error when the fluid is still; however, the calibration done under one circumstance may lose efficacy due to the environment dependence of the zero drift error [7]. Borg presented a "current source" to drive the transmit transducer which could achieve similar impedance symmetry of transmitting and receiving circuits [8]. In order to get high-power transmission, the current source is not an appropriate choice compared with the voltage source.

In a clamp-on type UFM, a vertically polarized shear wave generated by a clamp-on transducer will be reflected successively in the pipe wall and then used for measuring the signal launch [9]. The remainder of the wave could transfer to the receiver through the wall directly and be treated as reference signal for the zero drift error calibration. However, as the wave should transit along several different materials including the transducers, the pipe wall, and the gas or liquid flow, the intensity of the received signal is very weak even if the stimulation pulses are as large as several hundred voltages. Thus, the accuracy of the clamp-on type UFM is relatively low, even though various advanced digital signal process methods have been adopted.

This chapter proposes a novel ultrasonic transducer module with an own reference sound channel for zero error calibration. It consists of two energy conversion terminals and an own reference sound channel which is employed for the transmission of the reference signal instead of the pipe wall. The reference signal will be affected by the performance of transducers and electric circuits while the measuring signal is affected by the flow status additionally. According to the difference of reference signal and measuring signal, the calibration can be done in time. As the module is made of material whose acoustic impedance is specially designed to match the terminals and flow media, the module has much better transmitting and receiving efficiencies than the clamp-on method, which provides possibilities for developing high accuracy and low power assumption UFMs. The construction, the zero elimination method, and the design theory of the module have been illustrated in this chapter. The design has been validated by simulation. A prototype has been developed with the preliminary results demonstrating an observable decrease of the zero drift error.

93.2 Method and Design Theory

93.2.1 Method

93.2.1.1 Construction

The basic structure of the proposed transducer module is shown in Fig. 93.1. A pair of energy conversion terminals which play the roles of the transmitter and receiver are placed on the angled end face of the own reference sound channel. The own reference sound channel is employed for the transmission of the reference signal. For flow measurement application, the module is inserted into the measuring pipe to ensure that its lower face is on the same surface with the inner surface of the measuring pipe and contacts with the flow media directly.

93.2.1.2 Zero Elimination Method

Here, we take the case that terminal A is used for signal transmitting and B for receiving as example. The reference sound channel is made of material whose acoustic impedance is specially designed to match the terminals and flow media. As shown in the above figure, the ultrasound emitted from terminal A is separated into two parts at the interface of the module and the flow media. One part, shown as red solid lines in the figure, is transferred into the media, reflected at the opposite pipe wall and thus received by terminal B with the generated signal defined as the measuring signal for flow rate determination. The rest, shown as red dashed lines, is reflected on the interface of the module and the flow media, and propagates along the reference channel (inside the module) to terminal B with the generated signal defined as the reference signal for zero drift acquisition.

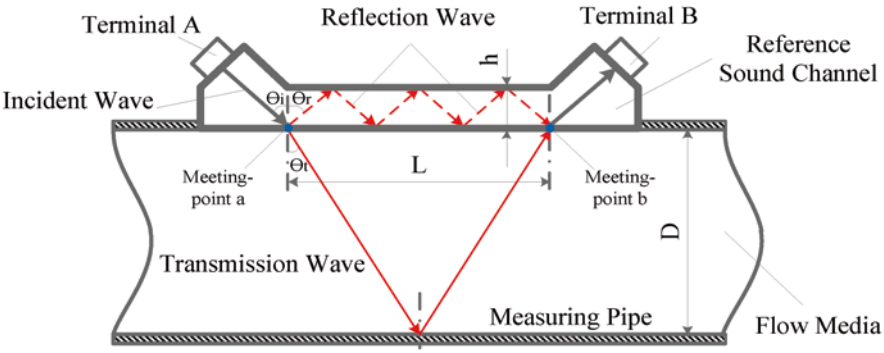


Fig. 93.1 Structure illustration of the module

The transit times of the measuring signal along upstream and downstream directions defined as t_{mAB} and t_{mBA}, respectively, can be formulated as:

$$t_{\mathrm{mAB}} = t_{\mathrm{eA}} + t_{\mathrm{fAB}} + t_{\mathrm{rB}} \tag{93.1}$$

$$t_{\mathrm{mBA}} = t_{\mathrm{eB}} + t_{\mathrm{fBA}} + t_{\mathrm{rA}} \tag{93.2}$$

where t_{eA} and t_{eB} are the delay times when terminal A and B are used as emitters, respectively; t_{rA} and t_{rB} are the delay times when terminal A and B are used as receivers, respectively; and t_{fAB} and t_{fBA} are the transit times in flow media along upstream and downstream directions, respectively. Δt_{m} can be calculated from Eqs. (93.1) and (93.2) as the transit time difference of the measuring signal. In the equation, Δt_{f} is the transit time difference caused by transferring in flow, based on which the flow rate of the media can be determined. Δt_0 is the delay time difference caused by terminals and electric circuits inconformity, which will change with environment conditions and thus cause a zero drift. The purpose of the novel design in this chapter is to eliminate Δt_0 in time.

$$\Delta t_{\mathrm{m}} = t_{\mathrm{mAB}} - t_{\mathrm{mBA}} = \Delta t_0 + \Delta t_{\mathrm{f}} \tag{93.3}$$

$$\Delta t_{\mathrm{f}} = t_{\mathrm{fAB}} - t_{\mathrm{fBA}} \tag{93.4}$$

$$\Delta t_0 = t_{\mathrm{eA}} + t_{\mathrm{rB}} - t_{\mathrm{eB}} - t_{\mathrm{rA}} \tag{93.5}$$

The transit times of the reference signal along different directions defined as t_{rAB} and t_{rBA}, respectively, can be formulated as:

$$t_{\mathrm{rAB}} = t_{\mathrm{eA}} + t_{\mathrm{cAB}} + t_{\mathrm{rB}} \tag{93.6}$$

$$t_{\mathrm{rBA}} = t_{\mathrm{eB}} + t_{\mathrm{cBA}} + t_{\mathrm{rA}} \tag{93.7}$$

where t_{cAB} and t_{cBA} are the transit times in the reference sound channel between the two meeting-points a and b of the transmission wave and reflection wave, respectively. Considering the reference sound channel is fixed without the influence of the flow status, $t_{\mathrm{cAB}} = t_{\mathrm{cBA}}$. Δt_{r} can be calculated from Eqs. (93.6) and (93.7) as the transit time difference of the reference signal.

$$\Delta t_{\mathrm{r}} = t_{\mathrm{rAB}} - t_{\mathrm{rBA}} = \Delta t_0 \tag{93.8}$$

According to Eqs. (93.3) and (93.8), Δt_{f} can be formulated as:

$$\Delta t_{\mathrm{f}} = \Delta t_{\mathrm{m}} - \Delta t_{\mathrm{r}} \tag{93.9}$$

Based on the two signals received, the zero drift error induced by environmental change can be determined from Eq. (93.9) and then calibrated.

93.2.2 Design Theory

To realize the above method, an important precondition is to ensure that the measuring signal and the reference signal arrived at the receiving terminal and both have enough intensity. Thus, it is critical to enhance and balance the energy transferred in both paths. Studies have been carried out including energy distribution design and sound path design based on acoustic theory. The former is aiming at distributing the ratio of energy transfers into different paths reasonably, while the latter is aiming at ensuring that both waves can be received vertically at the receiving terminal and have enough time intervals for the signal process.

93.2.2.1 Energy Distribution Design

Considerplane wave incidents on the interface of the reference sound channel and flow media. Assume that the wave only transfers in a 2D plane and ignore the time factor for the sake of simplicity. The sound pressure reflection coefficient E_r and the transmission coefficient E_t are [10]:

$$E_r \equiv \frac{p_{0r}(\omega)}{p_{0i}(\omega)} = \frac{m\cos\theta_i - \sqrt{n^2 - \sin^2\theta_i}}{m\cos\theta_i + n\cos\theta_t} \tag{93.10}$$

$$E_t \equiv \frac{p_{0t}(\omega)}{p_{0i}(\omega)} = \frac{2m\cos\theta_i}{m\cos\theta_i + \sqrt{n^2 - \sin^2\theta_i}} \tag{93.11}$$

$$m \equiv \rho_1 / \rho_0 \tag{93.12}$$

$$n \equiv c_0 / c_1 \tag{93.13}$$

where $p_{0r}(\omega)$, $p_{0t}(\omega)$, and $p_{0i}(\omega)$ are the sound pressure of the reflection wave, transmission wave, and incident wave, respectively; ρ_1 and ρ_0 are the density of the reference sound channel and flow media, respectively; c_1 and c_0 are the sound speed in the reference sound channel and flow media, respectively; θ_i is the incident angle, θ_r is the reflection angle, and θ_t is the transmission angle. According to Snell's law:

$$\theta_i = \theta_r \tag{93.14}$$

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{c_0}{c_1} \tag{93.15}$$

For the system shown in Fig. 93.1, the sound pressure reflection coefficient E_r, transmission coefficient E_t, reflection angle θ_r, and transmission angle θ_t can be determined by Eqs. (93.10) to (93.15) when the other parameters are decided. In order to achieve the most optimal voltage ratio of the measuring signal and reference signal, the incident angle θ_i and the material of the reference sound channel should be chosen appropriately.

93.2.2.2 Sound Paths Design

Based on the acoustic symmetry and the structure symmetry, the reflection wave and the transmission wave will arrive at another meeting-point b upon separation from the meeting-point a while the horizontal propagation distance of the two waves between meeting-points a and b remains the same, which can be formulated as:

$$kh \tan\theta_i = L \tag{93.16}$$

$$\frac{c_1}{c_0} \sin\theta_i = \frac{L}{\sqrt{4D^2 + L^2}} \tag{93.17}$$

where k is the times of reflection of the reflection wave in the reference sound channel, h is the thickness of the reference sound channel, L is the horizontal distance between meeting-points a and b, and D is the diameter of the pipe.

Furthermore, in order to keep the two received signals distinguishable while the frequencies remain the same, the constraint equation can be formulated as:

$$\frac{Lc_0}{c_1^2 \sin\theta_i} - \frac{kh}{c_0 \cos\theta_i} > iT_0 \tag{93.18}$$

where i is the number of excitation pulses, and T_0 is the period.

93.3 Simulation Validation and Prototype Development

A virtual transducer module has been designed and validated by numerical simulation. In addition, the actual prototype has been developed.

93.3.1 Design of the Virtual Transducer Module

To realize the module proposed above, the design of energy conversion terminals is a primary foundation. The traditional ultrasonic transducer used for flow measurements usually consists of the piezoelectric ceramic vibrator, the matching layer, and the backing layer in general. The piezoelectric ceramic vibrator is employed for the

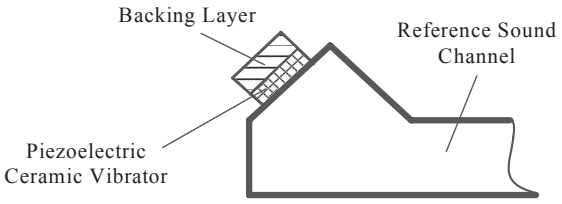


Fig. 93.2 Structural illustration of the energy conversion terminal

conversion of electric energy and acoustic energy. The matching layer is used for acoustic matching of piezoelectric ceramic vibrator and flow media. The backing layer is introduced to reduce the energy loss through the backing of the piezoelectric ceramic vibrator.

In the novel module, we integrate the normal matching layer into the reference sound channel since the material used remains the same. The energy conversion terminals consist only of a piezoelectric ceramic vibrator and a backing layer, which are introduced and fixed onto the reference sound channel. The model developed has ignored the shielded cable for simplification, as shown in Fig. 93.2.

Assume that the flow media is water without loss of generality. The piezoelectric ceramic vibrator is made of PZT-5 material with a resonance frequency equal to 1 MHz. The default parameters of water and PZT-5 material are selected as the simulation software has furnished. The backing layer is made up of the compound of epoxy resin and tungsten powder which is similar to the reference sound channel. However, in order to achieve the high impedance characteristic, the proportion of tungsten powder is elevated for the backing layer. Referring to the common matching layer, the proportion of different components for the reference sound channel is determined by the desired acoustic impedance $Z_{\mathbf{r}}$, which is given as:

$$Z_{\mathrm{r}} = \sqrt{Z_{\mathrm{p}} \cdot Z_{\mathrm{f}}} \tag{93.19}$$

where $Z_{\mathbf{p}}$ and $Z_{\mathbf{f}}$ are the acoustic impedance of the piezoelectric ceramic vibrator and flow media, respectively, which are related to the density and sound velocity of corresponding materials. The parameters of the reference sound channel and the backing layer are shown in Table 93.1.

According to the parameters shown in Table 93.1, the incident angle θ_{i} is selected to be 60° based on Eqs. (93.10) to (93.13); therefore, the reflection angle θ_{r} and the transmission angle θ_{t} are equal to 60° and 39.9°, respectively, based on Eqs. (93.14) and (93.15). The main sizes of the whole module are determined from Eqs. (93.16) to (93.18) as shown in Table 93.2.

Table 93.1 Parameters of the reference sound channel and the backing layer

Parameters	Poisson ratio	Density (kg/m^3)	Sound velocity (m/s)	Young modulus (GPa)
Reference sound channel	0.386	1721.6	2026.0	3.642
Backing layer	0.376	2779.7	1998.4	6.074

Table 93.2 Main sizes of the whole module, $k=3$

Parameters	L	h	D
Value (mm)	150	8	50

93.3.2 Numerical Validation

Based on the virtual transducer module developed above, we employ the acoustics module of COMSOL Multiphysics analysis software for numerical validation [11]. Considering that the variation of the acoustic intensity along the normal direction of paper is small, ignore the variation and build the model with the 2D method, as shown in Fig. 93.3. The acoustic-structure boundary of flow media is set as plane wave radiation. The maximum element size of flow media, the piezoelectric ceramic vibrator, the reference sound channel, and the backing layer is 0.33, 2, 4, and 4 mm, respectively.

The received signals (i.e., measuring signal and reference signal) in the static state are calculated. In order to elaborate the relationship of the two signals, the normalization method, i.e., regarding the measuring signal as the standard signal, is introduced to process the calculation results as shown in Fig. 93.4. According to Eq. (93.7), the front wave packet is the reference signal and the second is the measuring signal. Observing from the figure, the amplitude of the reference signal nearly reaches 50% of the measuring signal, which is large enough for signal processing. It is a balanced result based on sound path design. If the two signals get too close to each other, the signal resolution will decrease and affect the determination of arrival time. On the other hand, the difference of signal amplitudes will increase while the two signals separate too far away, which will lead to more difficulties in terms of signal acquisition. The optimization of sound paths will be a main point in future studies.

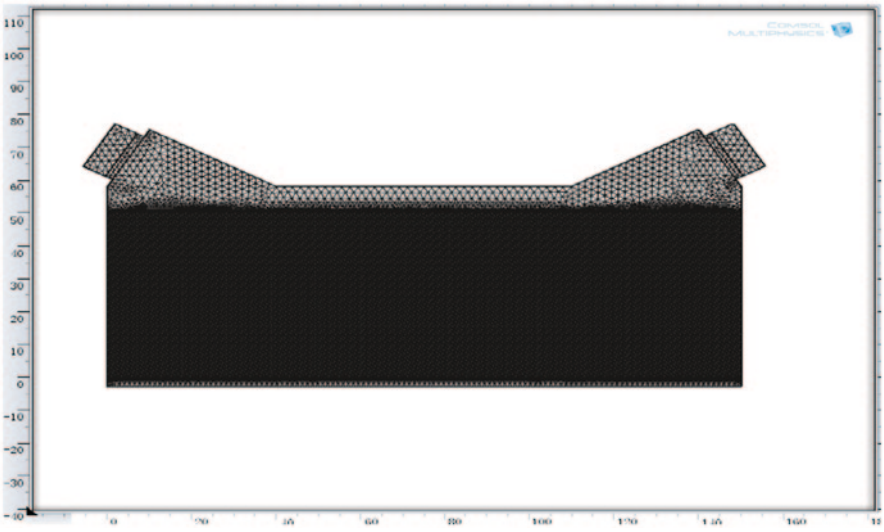


Fig. 93.3 Meshing results of the finite element model

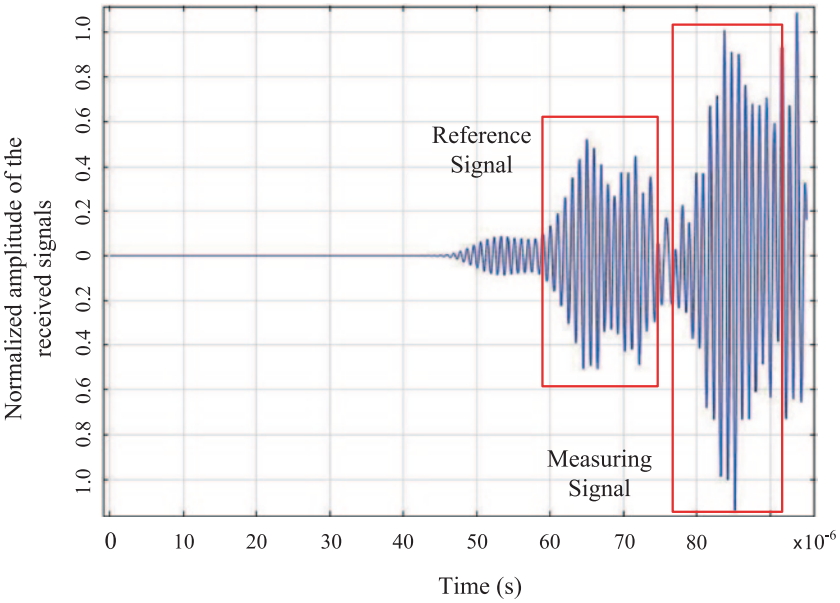


Fig. 93.4 The calculated received signals in static state

93.3.3 Prototype Development

According to the design scheme proposed in Fig. 93.1, we have developed a prototype as shown in Fig. 93.5.

The device shown in Fig. 93.5 is settled in the static open water field. A series of preliminary tests have been implemented under different environment conditions by adjusting the water temperature. The zero drift error is decreased for almost one order of magnitude compared with traditional flow meters. Further tests will be done when the customized pipe is processed.

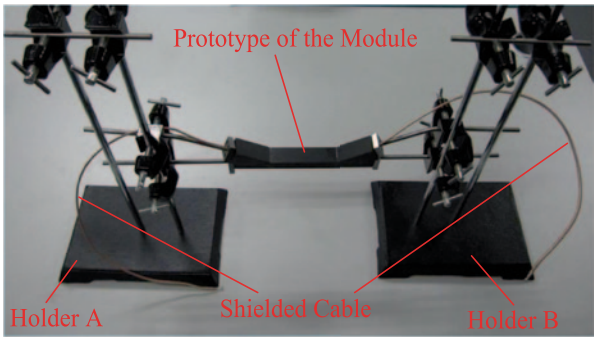


Fig. 93.5 Prototype developed and its holders for zero drift testing

93.4 Conclusion

We have proposed a novel ultrasonic transducer module with an own reference sound channel for flow measurement. According to the difference of measuring signal and the reference signal, the zero drift error can be calibrated. The virtual transducer module is introduced and the calculation results show that a reference signal nearly reaches 50 % of the amplitude of the measuring signal based on the special design of materials and construction, which enables the following signal processing. A prototype is developed and the preliminary tests demonstrate a decrease of the zero drift error for almost one order of magnitude compared with traditional flow meters. Based on the results obtained, the feasibility of the module is testified. The future work will focus on the implementation of the prototype installed into a practical UFM.

Acknowledgments The authors are grateful for the financial support from the National High Technology Research and Development Program (863 program) of China (No. 2012AA040605), the National Natural Science Foundation of China (No. 51105329), and the special support from the Postdoctoral Science Foundation of China (No. 2012T50509).

References

1. Marco D, Mauro C, Matteo D. Measurement of high-pressure natural gas flow using ultrasonic flowmeters. J Meas. 1997;20(2):75–89.
2. Bo Y, Li C, Yupin L. Forced oscillation to reduce zero flow error and thermal drift for non-reciprocal operating liquid ultrasonic flow meters. J Flow Meas Instrum. 2011;22(4):257–64.
3. Bo Y, Li C. Electronic circuit design for reciprocal operation of transit-time ultrasonic flow meters. J Flow Meas Instrum. 2013;32:5–13.
4. Lunde P, Vestrheim M, Reidar B, et al. Reciprocal operation of ultrasonic flow meters: criteria and applications. In: Marjorie Y, Editor. IEEE Ultrasonics Symposium, New York, U.S.A, 28–31 Oct 2007. Piscataway: IEEE; 2007. p. 381–6.
5. Willatzen M, Kamath H. Nonlinearities in ultrasonic flow measurement. J Flow Meas Instrum. 2008;19(2):79–84.
6. Yong C, Yiyong H, Xiaoqian C. Acoustic propagation in viscous fluid with uniform flow and a novel design methodology for ultrasonic flow meter. J Ultrason. 2013;53(2):595–606.
7. Coutard F, Tisserand E, Schweitzer P. The temperature influence on the piezoelectric transducer noise, measurements and modeling. In: Marjorie Y, Editor. IEEE Ultrasonics Symposium, Rotterdam, Netherlands, 19–21 Sept 2006. Piscataway: IEEE; 2005. p. 1652–5.
8. Borg J, Johansson J, van Deventer J, et al. Reciprocal operation of ultrasonic transducers: experimental results. In: Marjorie Y, Editor. IEEE Ultrasonics Symposium, Vancouver, Canada, 2–6 Oct 2006. Piscataway: IEEE; 2006. pp. 1013–6.
9. Xiaolei A, Lexington MA. Clamp-on steam/gas flow meter. USA: Patent 6626049; 2003.
10. Jianchun C. Basic principle of acoustics. 1st edn. Beijing: Science Press; 2012. p. 49–68.
11. Zhi Z. Device simulation and anisotropic structural design of the piezoelectric transducer based on the finite element method. MSc Thesis, Changchun: Jilin University, China; 2012 (In Chinese).

Part VIII
Control and Optimization

Chapter 94
The Lateral Channel Active Disturbance Rejection Controller Design of Small Air-to-Ground Missile for Helicopter

Yanpeng Hu, Shu Li and Hengxuan Gao

Abstract This chapter deals with the active disturbance rejection control (ADRC) method to the lateral control of small air-to-ground missiles for helicopters. The main difficulties that arise in lateral control design of small missiles are related to the highly uncertain, nonlinear aerodynamic forces and force movements, thus it is difficult to obtain an accurate dynamic model. Besides, the huge variations of dynamic pressure make controller parameters, and are difficult to adapt to the whole flight envelope. To handle them, the design of the controller is based on ADRC, a nonlinear control design approach which is based on the extended state observer (ESO), and can estimate and compensate the uncertainty of the system, including structural errors, random, and other system errors. Six-degree freedom equations are established for the lateral channel, the equations are linearization with small disturbance theory, then the aerodynamic derivatives are obtained through wind tunnel test, the dynamic model is made and the aerodynamic features are analyzed. One-order ESO equations of lateral channel are deduced, the controller of the lateral channel which is based on the ESO is designed and parameters are adjusted, and finally, semi-physical simulations are made. Through the comparison with classical proportion integral differential (PID) controller, the controller based on ADRC has strong robustness and it could track control command rapidly and accurately.

Keywords Small air-to-ground missile · Extended state observer · Active disturbances rejection control · Dynamics · Controller · Simulation

94.1 Introduction

At present, the majority of flight control systems in engineering are designed using the classical frequency domain method or root locus method. The classical control methods are based on the precise dynamic model. It demands the accurate parameters of aerodynamics throughout the flight envelope in order to have a high quality

Y. Hu (✉) · S. Li · H. Gao
BeiHang University, 100000 Beijing, China
e-mail: davidlove2008@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_94

control. For small diameter air-to-ground missiles, it is difficult to make an accurate model, and it is determined by the aerodynamic forces and moments that have an effect on the missile, which depends on flight altitude, velocity, attack angle, side-slip angle, rudders, etc. The variation of state causes the aerodynamic characteristics to show nonlinear and time-varying during the flight envelope, which makes it difficult for the linear controller to satisfy the accuracy.

This chapter adopts the active disturbance rejection control (ADRC) method which can estimate unknown dynamic model and external unknown disturbances, and compensate these unknown factors to controller design. The ADRC method realizes the feedback linearization of the dynamic system, and uses the nonlinear configuration consisting of the nonlinear feedback control law to improve the control performance of the closed-loop system [1].

94.2 The Controller Design

94.2.1 The Extended State Observer

In Fig. 94.1, b is a parameter related to the control amount. For block diagram, it can establish an extended state observer (ESO) (94.1) (94.2) [2],

$$\begin{cases} x_1 = x_2 \\ \dot{x}_2 = x_3 + bu \\ \dot{x}_3 = w(t) \\ y = x_1 \end{cases} \tag{94.1}$$

$$\begin{cases} e_1 = z_1 - y \\ \dot{z}_1 = z_2 - \beta_1 e_1 \\ \dot{z}_2 = z_3 - \beta_2 g_2(e_1) + bu \\ \dot{z}_3 = -\beta_3 g_2(e_1) \end{cases} \tag{94.2}$$

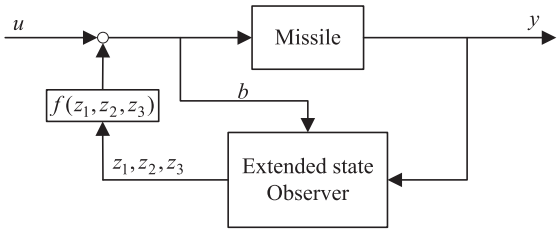


Fig. 94.1 Structure of the VSC based on ESO

Only through selecting suitable $\beta_1, \beta_2, \beta_3$ could get state estimate variables $z_1(t)$, $z_2(t)$ and $z_3(t)$, $f(x_1(t), x_2(t), t, w(t))$ of state $x_1(t)$ and $x_2(t)$ and the extended state variables. Whether f is continuous or not, as long as it is bounded and b is known, the state and extended state can be estimated well at real time, so the ESO is independent with the specific description of the transfer function [3-5].

94.2.2 The Control Characteristics of Small Missiles

The small missile has the features of small mass and small movement of inertia, which have a serious impact on the flight control. The variation of mass influences the transfer coefficient of transfer function, while the moment of inertia influences the damping of the missile. These characteristics make the missile sensitive to aerodynamics, especially m_x^{β}, which determines the control effects of the roll channel directly. Because of the small moments of inertia, the control performance of the roll channel is affected obviously by the disturbance, and less disturbance would produce an angle error that makes the roll channel difficult to control. In addition, because of serious varieties of flight dynamic pressure, and height variation in full envelope process, it has high requirements to the adaptability of parameters, and it also has high demands to the anti-disturbance of constant wind and wind shear.

94.2.3 The Establishment of Lateral Model

The lateral motion equations are, [6], (94.3, 94.4),

$$\begin{bmatrix} \dot{\Delta\omega}_x \\ \dot{\Delta\omega}_y \\ \dot{\Delta\beta} \\ \dot{\Delta\gamma} \end{bmatrix} = N \begin{bmatrix} \Delta\omega_x \\ \Delta\omega_y \\ \Delta\beta \\ \Delta\gamma \end{bmatrix} + \begin{bmatrix} b_{17} \\ 0 \\ 0 \\ 0 \end{bmatrix} \Delta\delta_x + \begin{bmatrix} b_{15} \\ b_{25} - b_{24}' b_{35} \\ -b_{35} \\ 0 \end{bmatrix} \Delta\delta_y + F \tag{94.3}$$

$$N = \begin{bmatrix} b_{11} & b_{12} & b_{14} & 0 \\ -b_{21} + b_{24}'\alpha & b_{22} - b_{24}' b_{32} - b_{24}'\alpha \tan\vartheta & b_{24} - b_{24}'(b_{34} - a_{33}) & -b_{36} b_{24}' \\ \alpha & -(\alpha \tan\vartheta + b_{32}) & -b_{34} + a_{33} & -b_{36} \\ 1 & -\tan\vartheta & 0 & 0 \end{bmatrix} \tag{94.4}$$

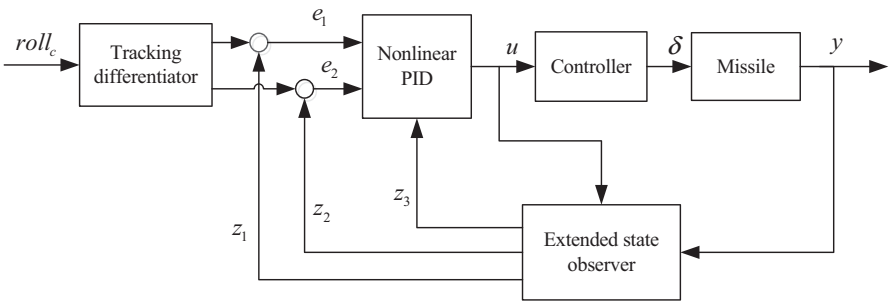


Fig. 94.2 Structure of the attitude controller based on ESO

The structure of the classical proportion integral differential (PID) controller is shown in Fig. 94.5. The angle that gyro measured and the command roll angle are the input signals of the controller and the control signals are inputs to rudders and then to the missile through the proportional, integral, and differential.

This chapter adopts the following attitude controller structure, which is based on the attitude angle outputs and control inputs of the system, and could utilize the extended observer to observe and compensate the disturbance of the system. The controller could realize the decoupling control of three-axis attitude, and compensate the disturbance torque during the flight envelope (Fig. 94.2).

It assumed the unknown disturbance system as (94.5, 94.6, 94.7)

$$\dot{x} = f(x, w, t) + bu(t) \tag{94.5}$$

$$f(x, w, t) = f_0(x, w, t) + f_1(x, w, t) \tag{94.6}$$

In the equation, $f_0(x, w, t)$ is a specific section of the model, $f_1(x, w, t)$ is an unknown section of the model, $w(t)$ is disturbance, and $u(t)$ is control information. The reduced order ESO is,

$$\begin{cases} \dot{z} = -cz - c^2 x - c\overline{b}u \\ \hat{x} = z + cx - f_0 \end{cases} \tag{94.7}$$

The output of ESO could estimate the unmodeled dynamics and external disturbances $f_1 + (b - \overline{b})u$ of real-time system, $\overline{b}$ is the estimation of controller input b, and b could be determined by the features of the system.

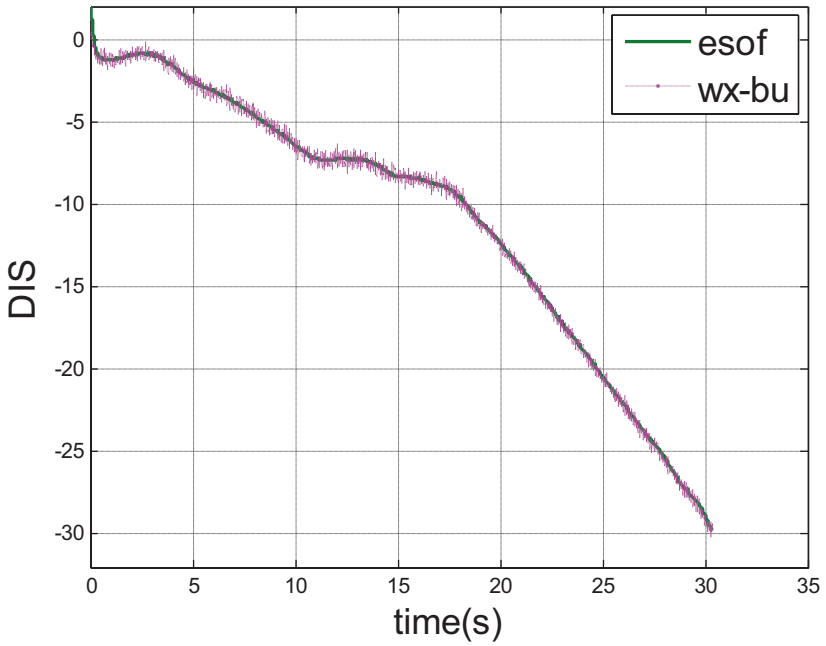


Fig. 94.3 The graph of *ESO* and *wx-bu*

94.3 Simulations

Figure 94.3 shows that the lines of the wx-bu and ESO are highly coincident. It shows that the ESO of the controller that is based on ESO could estimate and compensate the disturbance of the system accurately.

Figure 94.4 shows the performance of the PID controller and ADRC controller tracking sine signal, it indicates that the line of PID has a nearly 4° overshoot at first while it was only 1.5° in ESO [7-2002]. It also can be seen that the tracking is smooth after 8 s; besides, at the last 8 s, the bandwidth becomes higher than we designed. However, it is smooth for almost all the processes. The results show that ESO has a higher performance of tracking and rapidity; it can obtain control more accurately and quickly than the PID controller.

Figure 94.5 shows the response of roll and rolling angular rate in a flight simulation. In the simulation, the disturbance, structural deviations, and noise are added. The result shows that ESO has a lower overshoot at first than PID, and it can control both roll and rolling angular rate at zero very well, while the roll and rolling angular rate of PID fluctuate at zero. The simulation results indicate that the ESO can track the demand more accurately and rapidly, and has higher robustness.

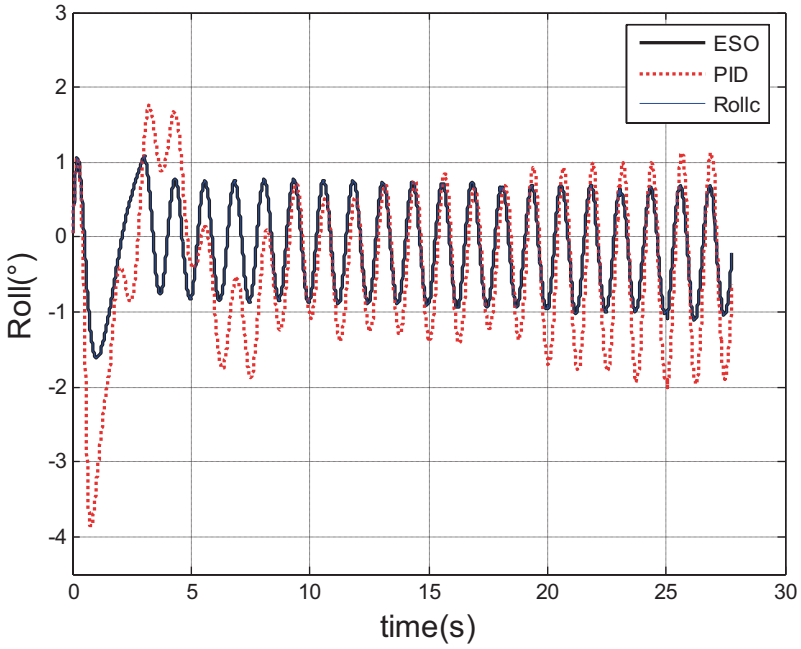


Fig. 94.4 The graph of tracking of sine command for *ESO* and *PID*

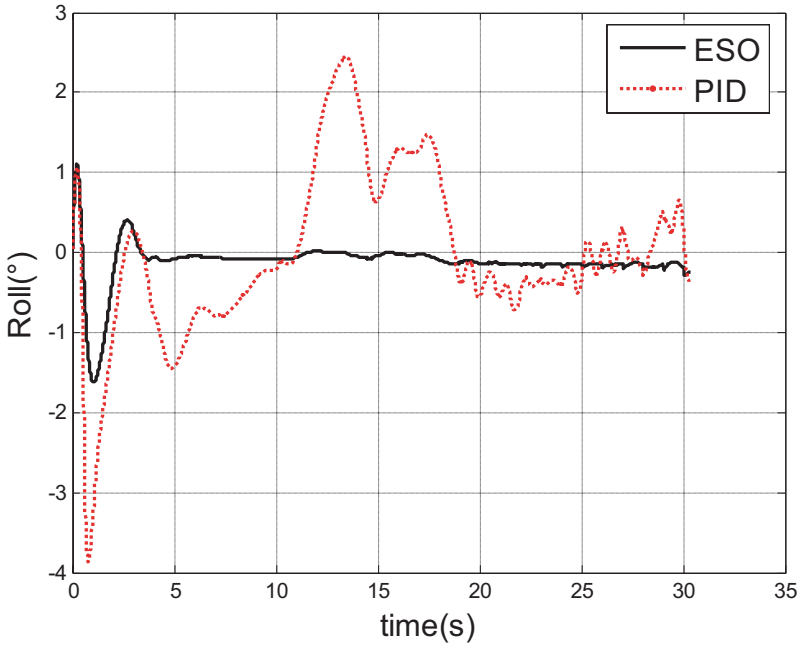


Fig. 94.5 The graph of comparison of roll response for *ESO* and *PID*

94.4 Conclusion

This chapter designed the extended state observer of roll channel, which was based on the ADRC method. The simulations were made for both PID control and ESO control, and the contrast of results showed that ESO could compensate the uncertainty accurately in the flight envelope, and have higher robustness and rapidity.

References

1. Chen J, Jin TX. Based on extended state observer variable structure control and its application. Transac Beijing Inst Technol. 2009;4:318–22.
2. Han JQ. From PID technology to ADRC technique. Control Eng China. 2002;9(3):13–8. (in Chinese).
3. Han JQ. Error-error principle—the basic idea of the anti-rejection control technology. Tsinghua University: Department of Thermal Engineering; 2005. (in Chinese).
4. Han JQ. Active disturbance rejection control technology. Beijing: National Defense Industry Press; 2008 pp. 197–07. (in Chinese).
5. Hu HY, Wen Y. The attitude controller design of damaged wing aircraft based on ADRC method. Flight Dyn. 2013;31(3):31–5. (in Chinese).
6. Qian XF, Lin RX, Zhao YN. Dynamics of missile. Beijing: Beijing Institute of Technology Press; 2000 p. 231–4. (in Chinese).
7. Sun CS, Tian J. Optimal sliding mode variable structure control of servo regulators with asymptotic observers. Proceedings of the 3thworld congress on intelligent control and automation. HeFei, P. R. China; 2000. pp. 3021–5.
8. Wu ST, Fei YH. Flight control system. Beijing: Bei Hang University Press; 2005. p. 218–38. (in Chinese).
9. Xiao YH. The reaching law for variable structure control of discrete time system based on attenuating control. Control Theory Appl 2002;19(2):450–52. (in Chinese).

Chapter 95
The Design of a Smart Car Based on a Kind of Dynamic Threshold Algorithm and Speed Control Algorithm

Qiyan Ke, Dashun Que and Wei Lu

Abstract Now smart cars have become one current researching hotspot in the field of intelligent control. This chapter designs and establishes a smart car system based on a kind of dynamic threshold algorithm and proportion integration differentiation (PID) control algorithm after researching the existing control algorithms. The core is to identify the black line on both sides of the track, the bimodal threshold, and make the servo response quickly. At last, the algorithm simulation, the system implementation, and algorithm performance evaluation of the smart car is completed. The results of the experiments reveal that the algorithm and the smart car system are effective.

Keywords Bimodal threshold algorithm · PID control algorithm · Smart car · Tracking

95.1 Introduction

The smart car model is the micrographics model of the future smart cars, it can run at high speed and avoid obstacles automatically with the introduction of a line mark [1]. The national university's the Freescale cup intelligent car competition is the background of this chapter. The hardware of the smart car system takes the 32 bit Freescale microprocessor MK60DN512ZVLQ10 as the control core, and the system has the power management module, the main controller module, path identification module, speed detection module, motor drive control module, and actuator control module [2].

Q. Ke (✉) · W. Lu
The School of Information Engineering, Wuhan University of Technology, 430070 Wuhan, China
e-mail: keqiyan@whut.edu.cn

D. Que
Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education, Wuhan University of Technology, 430070 Wuhan, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_95

In this chapter, a new control algorithm will be mainly discussed. The new control algorithm is based on the bimodal threshold algorithm and the proportion integration differentiation (PID) control algorithm. The width of the track is not less than 45 cm. And the middle part of the track is white KT boards, on both sides of the track are two black lines. Outside the track is blue cloth. They will form two gray value peaks near the two black lines. Selecting the gray value of the peaks as a threshold will give a reasonable image segmentation. And this is just the bimodal threshold algorithm. PID control, is proportionally the deviation (P), integral (I), and differential (D) control of PID controller. In general, the location of the peaks and valleys is difficult to confirm, because the light is not stable. The statistical method to calculate the threshold will be explained in this chapter. The parameters of the algorithm are variable to make the smart car adapt to the environment under different light sources and different types of the track.

95.2 The Design of the Tracking Algorithm

The design of the system uses the linear CCD TSL1401 as the sensor. The output is nominally 0 V for no light input and 2 V for a nominal full-scale output. The linear CCD module will transform sampling voltages to gray values 0–255. Thus, first, the design of the system uses these gray values to calculate a threshold; then, changes gray values to the binary data 0 and 1 based on the threshold.

In general, segmentation algorithms can be divided into static threshold segmentation and dynamic threshold segmentation [3]. The traditional static threshold algorithm just compares collected gray value with the threshold and judges white line or black line. This method is simple, however, it has to meet the conditions that there cannot be other light interferences and the threshold has to be changed by people in different environments. After the study, it was found that the static threshold algorithm will be failure in the strong light environment. And the dynamic threshold algorithm can adapt to different light environments, changing the region segmentation, and make up the shortfalls of the static threshold algorithm. Many people have put forward various dynamic threshold algorithms. However, these algorithms are too limited or too complex, and cannot be applied to the smart car well. Accordingly, the design of this system developed a dynamic threshold algorithm to apply to the smart car.

Based on the characteristics of the track, we use a statistical method to calculate the two places, namely the bimodal threshold algorithm [4]. From the two positions, the threshold can be calculated. And the two formed peaks of the track are shown in Fig. 95.1. The bimodal threshold algorithm is very suitable for the characteristics of the track, and the threshold could be calculated easily compared to other dynamic threshold algorithms. It is capable of being applied to the smart car perfectly.

According to the effective range of array data (*GValueMin, GValueMax*), the data are divided into *N* space. The number of data in every space is NUM:

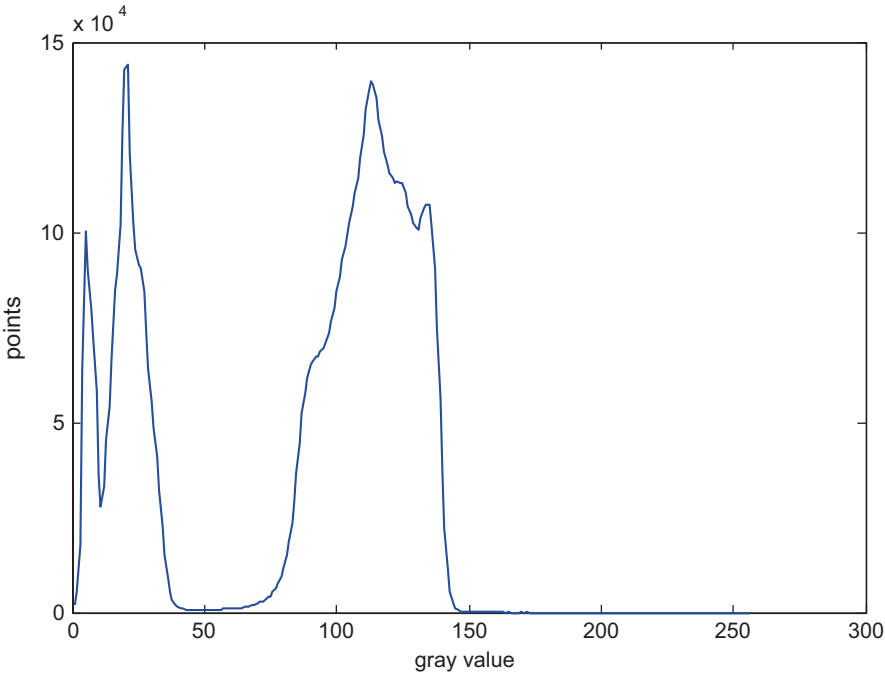


Fig. 95.1 The two formed peaks of the track

$$NUM = \frac{GValueMax\text{-}GValueMin}{N} \tag{95.1}$$

Count the number of the array data in these intervals (*GValueMin, GValueMin + NUM*), (*GValueMin + NUM, GValueMin + 2*NUM*)…(*GValueMin* + (*N*−*1*)**NUM, GValueMax*). According to the two peaks in the histogram corresponding to the horizontal "*Hleft*" and "*Hright*," the threshold could be calculated.

$$Threshold = \frac{Hleft + Hright}{2} * NUM + GValueMin \tag{95.2}$$

While the threshold has been calculated, the data can be thresholding.

$$CCDValue = \begin{cases} 0(GValue<Threshold) \\ 1(GValue>=Threshold) \end{cases} \tag{95.3}$$

Finding the data 0–1 on the edge and 1 to 0 falling edge can identify left edge (*LeftEdge*) and right edge (*RightEdge*). There is a formula for calculating the error, which will exclude inaccurate data. Steering deviation formula is as follows:

$$Error = \frac{LeftEdge + RightEdge}{2} - 64 \tag{95.4}$$

In the next part of this chapter, the PID control algorithm applicable to the servo will be expounded.

95.3 The PID Control Algorithm for the Servo

For the closed-loop control of the servo, the design used is the position type PID. This algorithm has good response speed, low overshoot, and low steady-state error [5]. The PID control algorithm block diagram is shown in Fig. 95.2.

$$U_k = K_p * e_k + K_i \sum_{j=0}^{k} e_k + K_d (e_k - e_{k-1}) \tag{95.5}$$

The actual tests showed that the two black lines center and the steering angle reference value are in a linear relationship. For experiments at low speed, set K_p to 0 when the smart car deviates from the black line in a range of rarity.

However, after increasing the speed up to 2.5 m/s, the smart car jittered seriously especially on the straights. A more intelligent PID control strategy has been developed based on these factors. Set K_i to zero, and set K_d a constant value K_p and the "Error" value are a quadratic function.

$$K_p = \frac{Error * Error}{2} + 1000 \tag{95.6}$$

By continuous debugging, a set of PID parameters was chosen to make the smart car trace smoothly and fast. The PID parameters tuning form is shown in Table 95.1 and the fitted curve is shown in Fig. 95.3.

The program flow chart is shown in Fig. 95.4. The smart car can run on the track smoothly based on the algorithm. The steering angle responses fast based on the position PID control algorithm, and the results are shown in Fig. 95.5. And from the experimental results, the servo responses were observed to be quick.

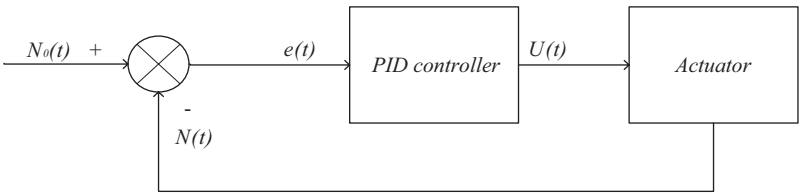


Fig. 95.2 PID control algorithm block diagram

Table 95.1 PID parameters tuning

The track type	Error	K_p	K_d	K_i	Results
The cross-shaped curve (turn left)	55	2512.5	4000	0	Run in the middle
The cross-shaped curve (turn right)	−54	2458	4000	0	The midline to the left
S-curve to 270° sharp curve(left)	70	3450	4000	0	Run in the middle
S-curve to 270° sharp curve(right)	−79	4120.5	4000	0	Run in the middle
The large S-curve(turn left)	46	2058	4000	0	Run in the middle
The large S-curve(turn right)	−42	1882	4000	0	Run in the middle
The U-shaped curve(turn left)	53	2404.5	4000	0	Cut the inside curve
The U-shaped curve(turn right)	−61	4721	4000	0	Cut the curve a little
Smooth curve to U-shaped curve(left)	63	2984.5	4000	0	Cut inside curve a little
Smooth curve to U-curve(right)	−71	3520.5	4000	0	Cut inside curve a little
The straight	0	1000	4000	0	Run in the middle

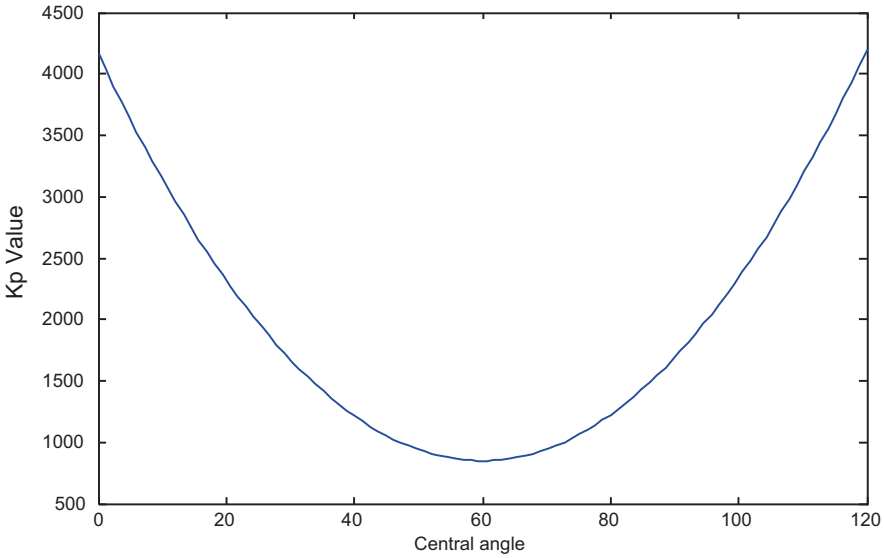


Fig. 95.3 The actual K_p value in different types of the track

95.4 The Main Traction Control Strategies

The main traction control strategies are mainly based on the speed of the smart car. A speed control algorithm which is also based on the PID control algorithm has been designed. The speed PID control algorithm takes after the turning PID control algorithm mentioned above, the difference is that the actuator is the motor but not the servo.

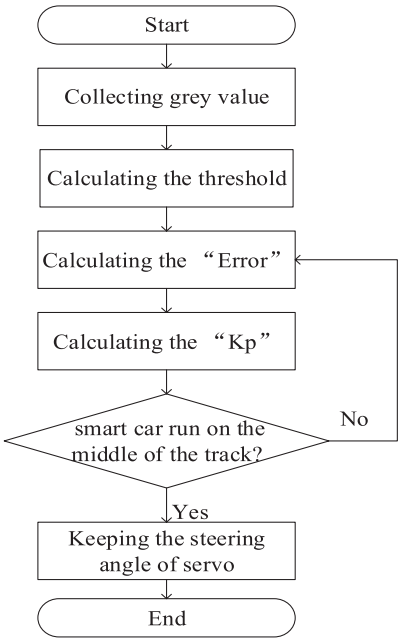


Fig. 95.4 Program flow chart of the PID control algorithm

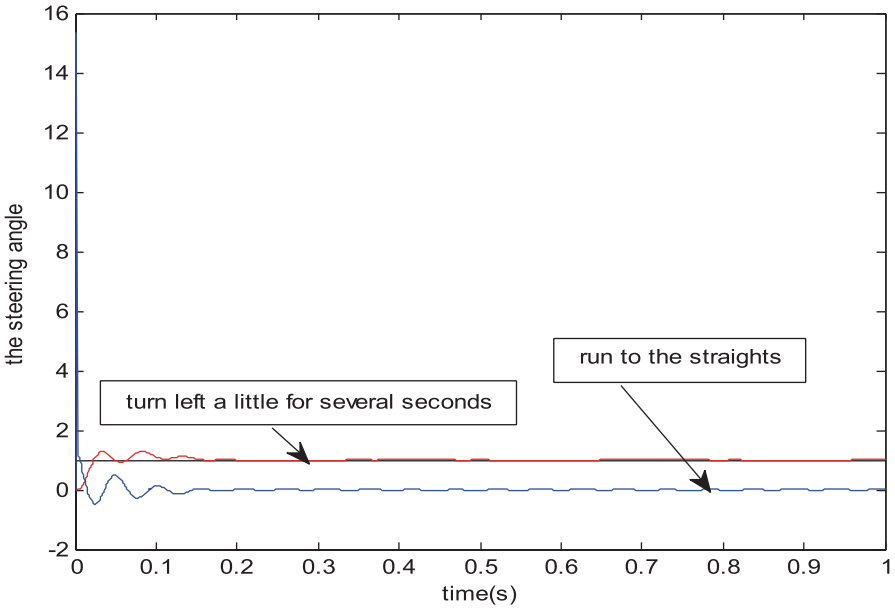


Fig. 95.5 The servo response of the steering angle

The basic control strategy is that the smart car accelerates on the straights and decelerates on the curves based on the speed PID control algorithm. At the same time, the smart car should adjust its position to make itself run on the middle line of the track based on the turning PID control algorithm for the servo.

About the cross-shaped curve, the smart car has to identify it. The both sides of the cross-shaped curve have no black lines, so if the track is whole white, keep the servo at the zero angle. The strategy of running out of the cross-shaped curve is like running into it, keeping the servo at the zero angle and letting the smart car run straight.

Code:

```
if(left_num == Window_Width_L&&right_num == Window_Width_R)
// Judge all black and all white cases, do consider all white
{
    if(left404==1||right404==1)
    Crossroad = 0;
    else Crossroad = 1;
}
```

About the mandatory turn crossing, the smart car has two TSL1401 linear CCD modules. The one identifies the regular tracks, and another identifies the mandatory turn crossing. When the second TSL1401 linear CCD module sees the Zebra crossing, letting the smart car turn 60° angle and decelerates. The second TSL1401 linear CCD module is also used to identify the roadblock. The servo of the smart car will turn a few degrees and then run straight when the second TSL1401 linear CCD module sees the whole black roadblock. The designed smart car is shown in Fig. 95.6.

Through the experiments, the smart car can speed up about 2.3 m/s on the curves, 2.7 m/s on the straights and run smoothly.

Fig. 95.6 Physical map of the smart car

95.5 Conclusion

This chapter adopts a kind of dynamic threshold algorithm based on the bimodal threshold algorithm and adopts the PID control algorithm to design a smart car control system. These two algorithms are jointly performed for tracking the line. By this method, the smart car can run on the track smoothly and fast. The theoretical analyses and the results obtained from the experiments reveal that the algorithm and the smart car system are effective.

References

1. Zhan H, Deng L, Xue S. Design of smart car speed control system based on fuzzy algorithm and PID algorithm. Power Electronics and Motion Control Conference (IPEMC), 7th International. IEEE conference publications; 2012. pp. 763–6.
2. Lin K, Huang T, Liu P. Design of the smart car control system based on adaptive PID algorithm. Conference Anthology, IEEE; 2013. pp. 1–4.
3. Du W, Tian X, Sun Y. A dynamic threshold edge-preserving smoothing segmentation algorithm for anterior chamber OCT images based on modified histogram. In: Image and Signal Processing (CISP), 4th International Congress on, IEEE; 2011. pp. 1123–6.
4. Patel SB, Patel PL, Porey PD. Threshold for initiation of motion of unimodal and bimodal sediments. Int J Sediment Res. 2013;28(1):24–33.
5. Jia Y. Steering control of intelligent vehicle based on the synthesis of fuzzy control arithmetic. Comput Knowl Technol. 2008;4(34):1877–8.

Chapter 96
Control Strategy of a Hybrid Active Filter Based on Output Voltage Correction

Yonghai Xu and Congxi Wen

Abstract A control strategy of a hybrid active filter based on output voltage correction is proposed. The active power filter (APF) is controlled as a voltage source. To improve the filtering performance, a sinusoidal signal is calculated as the APF output voltage correction signal on the basis of proportional control. The amplitude and phase of the correction signal are calculated based on those of system residual harmonic current. The sliding window discrete-fourier-transform (DFT) method is used to detect the amplitude and phase of the system residual harmonic current. A simple method is given to determine the parameters of the control strategy. The simulation results show that selection of the parameter is simple and effective, the APF responds fast to load harmonic current changes and performs well under background harmonic with the proposed method.

Keywords Hybrid active power filter · Proportional control · APF output voltage correction · System residual harmonic current · Background harmonic

96.1 Introduction

Installing filtering devices is the most common way to reduce harmonics, and the hybrid power filter (HPF) combines the respective advantages of passive and active filters with good prospects.

An HPF structure is proposed where a passive filter is in parallel with the load, and an active filter is in parallel with the reactor of the passive filter [1]. System resonance will be inhibited and the capacity of the active filter will be small with this kind of hybrid active power filter (APF) [2, 3]. For this hybrid filter structure, proportional control is applied with multiplied harmonic current as the reference value of the APF output voltage [1]. Multiplied harmonic current is used as the reference value of the APF output current [2]. But there exist conflicts between the filter performance and the stability of the system in both methods. Generalized integral

C. Wen (✉) · Y. Xu
North China Electric Power University, 102206 Beijing, China
e-mail: wcxmvp@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_96

control method is proposed to eliminate harmonic selectively which improves the filtration rate, but the dynamic response is slow [4]. A model-based control method is proposed, but the parameters are difficult to determine [5].

A control strategy based on output voltage correction for the HPF mentioned above is proposed in this chapter. On the basis of proportional control, a sinusoidal signal is calculated as the APF output voltage correction signal. Simulation shows that the proposed method is simple and effective to select the parameter and performs well even under background harmonic with high filtration rate.

96.2 The Analysis of Proportion Control

The hybrid filter structure is shown in Fig. 96.1a. The h^{th} order harmonic current of system $\vec{I}_{\text{Sh}}$ is divided into $\vec{I}_{\text{Sh0}}$ and $\vec{I}_{\text{ShAPF}}$, which are generated by the harmonic source and APF, respectively. $\vec{I}_{\text{Sh0}}$ and $\vec{I}_{\text{ShAPF}}$ are shown in Fig. 96.1b and Eq. (96.1).

$$\vec{I}_{\text{Sh}} = \vec{I}_{\text{Sh0}} + \vec{I}_{\text{ShAPF}} \tag{96.1}$$

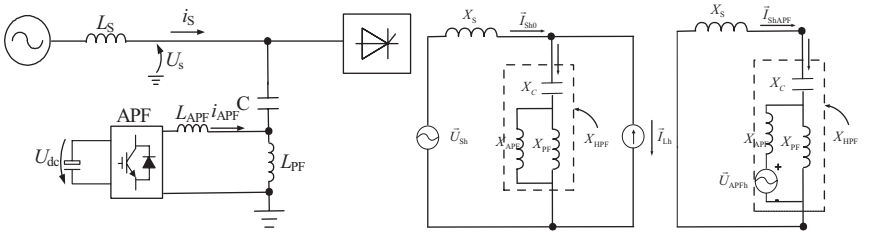
where

$$\vec{I}_{\text{ShAPF}} = -\frac{\vec{U}_{\text{APFh}} X_{\text{PF}}}{X_{\text{APF}}(X_{\text{PF}} + X_C + X_{\text{S}}) + X_{\text{PF}}(X_C + X_{\text{S}})} \tag{96.2}$$

The proportional control method is to control the APF output harmonic voltage as k times of the system harmonic current as Eq. (96.3) shows [1].

$$\vec{U}_{\text{APFh}} = k\vec{I}_{\text{Sh}} \tag{96.3}$$

Ignoring the resistance of the system and assuming the impedance between $\vec{I}_{\text{ShAPF}}$ and $\vec{U}_{\text{APFh}}$ is capacitive, Eq. (96.4) is deduced according to Eqs. (96.2) and (96.3).



Structure of HPF — The *h* order harmonic equivalent circuit of HPF

a **b**

Fig. 96.1 Structure and equivalent circuit of HPF. **a** Structure of HPF. **b** The *h* order harmonic equivalent circuit of HPF

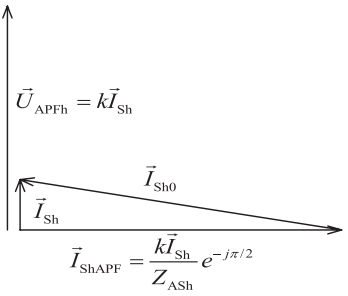


Fig. 96.2 The vector diagram of voltage and current

$$\vec{I}_{\mathrm{ShAPF}} = \frac{k\vec{I}_{\mathrm{Sh}}}{Z_{\mathrm{ASh}}} e^{-j\pi/2} \tag{96.4}$$

where $Z_{\mathrm{ASh}} = \left| \frac{X_{\mathrm{APF}}(X_{\mathrm{PF}} + X_C + X_{\mathrm{S}}) + X_{\mathrm{PF}}(X_C + X_{\mathrm{S}})}{X_{\mathrm{PF}}} \right|$.

According to Eq. (96.1), Eq. (96.5) can be obtained:

$$\vec{I}_{\mathrm{Sh0}} = \vec{I}_{\mathrm{Sh}} - \vec{I}_{\mathrm{ShAPF}} \tag{96.5}$$

According to Eqs. (96.3) to (96.5), the relationship among the vectors of $\vec{U}_{APFh}, \vec{I}_{\mathrm{Sh}}, = \vec{I}_{\mathrm{ShAPF}}, -\vec{I}_{\mathrm{Sh0}}$ is shown in Fig. 96.2 with counterclockwise as the positive direction.According to Fig. 96.2, the residual system harmonic current satisfies Eq. (96.6) under the proportional control method.

$$\vec{I}_{\mathrm{Sh}} = \frac{Z_{\mathrm{ASh}}}{\sqrt{k^2 + Z_{\mathrm{ASh}}^2}} \vec{I}_{\mathrm{Sh0}} e^{-j\arctan\frac{k}{Z_{\mathrm{ASh}}}} \tag{96.6}$$

Similarly, if the impedance between $\vec{I}_{\mathrm{ShAPF}}$ and $\vec{U}_{\mathrm{APFh}}$ is inductive, residual system harmonic current satisfies Eq. (96.7). The following analysis is based on Eq. (96.6):

$$\vec{I}_{\mathrm{Sh}} = \frac{Z_{\mathrm{ASh}}}{\sqrt{k^2 + Z_{\mathrm{ASh}}^2}} \vec{I}_{\mathrm{Sh0}} e^{j\arctan\frac{k}{Z_{\mathrm{ASh}}}} \tag{96.7}$$

96.3 Control Strategy of the Hybrid Active Filter Based on Output Voltage Correction

96.3.1 Principle of Output Voltage Correction

On the basis of proportional control, a sinusoidal signal is calculated as the APF output voltage correction signal. $\vec{U}_{\mathrm{APFh1}}$ represents the proportional control signal,

and $\vec{U}_{\mathrm{APFh2}}$ represents the correction signal. Then system harmonic currents generated by $\vec{U}_{\mathrm{APFh1}}$ and $\vec{U}_{\mathrm{APFh2}}$ are shown as follows:

$$\vec{I}_{\mathrm{ShAPF1}} = \frac{\vec{U}_{\mathrm{APFh1}}}{Z_{\mathrm{ASh}}} e^{-j\pi/2} \qquad \vec{I}_{\mathrm{ShAPF2}} = \frac{\vec{U}_{\mathrm{APFh2}}}{Z_{\mathrm{ASh}}} e^{-j\pi/2} \tag{96.8}$$

Supposing $\vec{I}'_{\mathrm{Sh0}} = \vec{I}_{\mathrm{Sh0}} + \vec{I}_{\mathrm{ShAPF2}}$, Eq. (96.1) can be transformed as Eq. (96.9):

$$\vec{I}_{\mathrm{Sh}} = \vec{I}'_{\mathrm{Sh0}} + \vec{I}_{\mathrm{ShAPF1}} \tag{96.9}$$

The proportional control signal changes as the system current harmonics changes, which regulates fast, and the correction signal is constant in a control cycle. The APF output voltage generated by the correction signal can be regarded as a stable sinusoidal voltage source. Therefore $\vec{I}_{\mathrm{Sh}}$ and $\vec{I}'_{\mathrm{Sh0}}$ satisfy the relationship in Eq. (96.6), which is shown in Eq. (96.10):

$$\vec{I}_{\mathrm{Sh}} = \frac{Z_{\mathrm{ASh}}}{\sqrt{k^2 + Z_{\mathrm{ASh}}^2}} \vec{I}'_{\mathrm{Sh0}} e^{-j\arctan\frac{k}{Z_{\mathrm{ASh}}}} \tag{96.10}$$

Making the time when APF is put into operation as moment 0, $\vec{U}_{\mathrm{APFh2}(i)}$, $\vec{I}_{\mathrm{ShAPF2}(i)}$, and $\vec{I}_{\mathrm{Sh}(i)}$ represent the APF output voltage correction signal, system harmonic current generated by $\vec{U}_{\mathrm{APFh2}(i)}$ and residual system harmonic current at moment i, respectively, making $\vec{I}'_{\mathrm{Sh0}(i)} = \vec{I}_{\mathrm{Sh0}} + \vec{I}_{\mathrm{ShAPF2}(i)}$. $\vec{I}_{\mathrm{Sh}(i)}$ is the result under the effect of $\vec{U}_{\mathrm{APFh2}(i-1)}$. Therefore Eq. (96.10) is converted into a discrete form as shown in Eq. (96.11):

$$\vec{I}_{\mathrm{Sh}(i)} = \frac{Z_{\mathrm{ASh}}}{\sqrt{k^2 + Z_{\mathrm{ASh}}^2}} \vec{I}'_{\mathrm{Sh0}(i-1)} e^{-j\arctan\frac{k}{Z_{\mathrm{ASh}}}} \tag{96.11}$$

Making $\vec{U}_{\mathrm{APFh2}(0)}=0$, $\vec{I}'_{\mathrm{Sh0}(0)} = \vec{I}_{\mathrm{Sh0}}$, $\vec{U}_{\mathrm{ShAPF2}}$ is adjusted according to Eq. (96.12) with $\vec{I}_{\mathrm{Sh}}$ as feedback:

$$\vec{U}_{\mathrm{APFh2}(i+1)} = \vec{U}_{\mathrm{APFh2}(i)} + m\vec{I}_{\mathrm{Sh}(i+1)} \tag{96.12}$$

where correction parameter m is a scalar constant. Equation (96.13) can be obtained by combining Eqs. (96.8), (96.11), and (96.12):

$$\vec{I}_{\mathrm{ShAPF2}(i)} = \frac{\vec{U}_{\mathrm{APFh2}(i)}}{Z_{\mathrm{ASh}}} e^{-j\pi/2} = \vec{I}_{\mathrm{ShAPF2}(i-1)} - \frac{m}{\sqrt{k^2 + Z_{\mathrm{ASh}}^2}} \vec{I}'_{\mathrm{Sh0}(i-1)} e^{j\left(\frac{\pi}{2} - \arctan\frac{k}{Z_{\mathrm{ASh}}}\right)} \tag{96.13}$$

Equation (96.14) is deduced by adding $\vec{i}_{\mathrm{Sh0}}$ to both sides of Eq. (96.13):

$$\begin{aligned}\vec{I}'_{\mathrm{Sh0}(i)} &= \vec{I}'_{\mathrm{Sh0}(i-1)} - \frac{m}{\sqrt{k^2+Z_{\mathrm{ASh}}^2}} e^{j\left(\frac{\pi}{2}-\arctan\frac{k}{Z_{\mathrm{ASh}}}\right)} \vec{I}'_{\mathrm{Sh0}(i-1)} \\ &= \left(1 - \frac{m}{\sqrt{k^2+Z_{\mathrm{ASh}}^2}} e^{j\left(\frac{\pi}{2}-\arctan\frac{k}{Z_{\mathrm{ASh}}}\right)}\right) \vec{I}'_{\mathrm{Sh0}(i-1)} = f\vec{I}'_{\mathrm{Sh0}(i-1)}\end{aligned} \tag{96.14}$$

Where $f = 1 - \frac{m}{\sqrt{k^2+Z_{\mathrm{ASh}}^2}} e^{j\left(\frac{\pi}{2}-\arctan\frac{k}{Z_{\mathrm{ASh}}}\right)}$. As long as $|f|<1$, $\vec{I}'_{\mathrm{Sh0}}$ will decrease with the increase of moment i, and residual system harmonic current $\vec{I}_{\mathrm{Sh}}$ will be reduced, eventually approaching zero.

96.3.2 Determination of Control Parameters

The magnitude of $|f|$ reflects the filtering performance mostly. If $|f|<1$, residual system harmonic current will be eliminated and the smaller $|f|$ is, the better the dynamic performance will be. Provided $m'=m/Z_{\mathrm{ASh}}$, $k'=k/Z_{\mathrm{ASh}}$, then

$$f = 1 - \frac{m'}{\sqrt{k'^2+1}} e^{j\left(\frac{\pi}{2}-\arctan k'\right)} \tag{96.15}$$

It is calculated that $|f| = \sqrt{(m'-k')^2+1} \Big/ \sqrt{k'^2+1}$. Therefore, the necessary and sufficient condition of $|f|<1$ is $0<m<2k$. Keeping k' unchanged, when $m'=k'$, $|f|$ is the minimum. If $m'=k'$, the greater k' is, the smaller $|f|$ will be. Therefore, from the view of filter performance, the bigger k the better, and it should be made that $m=k$.

The stability of the system is mainly affected by the proportional control parameter k. As long as $|f|<1$ the residual system harmonic current will eventually approach zero. Therefore, when choosing k, stability can be considered first to meet the requirements, without having to overemphasize the attenuation factor of harmonic which can cause instability. The stability margin can be calculated with the Nyquist method to determine if k is suitable [6].

96.3.3 Calculation of the Correction Signal

According to Eq. (96.12), the phase θ_{hi} and amplitude A_{hi} of $\vec{U}_{\mathrm{APFh2}(i)}$ can be calculated with Eq. (96.16):

$$A_{hi} = \sqrt{x_{hi}{}^2 + y_{hi}{}^2} \qquad \theta_{hi} = \begin{cases} \arctan(y_{hi}/x_{hi}) & x_{hi}>0 \\ \arctan(y_{hi}/x_{hi})+\pi & x_{hi}<0 \end{cases} \tag{96.16}$$

where x_{hi} and y_{hi} are expressed in Eq. (96.17):

$$\begin{cases} x_{hi} = A_{i-1}\cos(\theta_{i-1}) + dA_i\cos(d\theta_i) \\ y_{hi} = A_{i-1}\sin(\theta_{i-1}) + dA_i\sin(d\theta_i) \end{cases} \quad \begin{cases} dA_i = ma_{hi} \\ d\theta_i = \varphi_{hi} \end{cases} \tag{96.17}$$

where a_{hi} and φ_{hi} denote the amplitude and phase of the h^{th} order residual system harmonic current $\vec{I}_{\text{Sh}}$, respectively.

Therefore, the amplitude and phase of residual system harmonic current need to be calculated to obtain the APF output voltage correction signal. The sliding window iterative discrete-fourier-transform (DFT) method can be used to detect any order of the harmonic current fast with a small amount of computation [7]. S_{ah} and S_{bh} are two variables in the process of calculating the h^{th} order of harmonic current using this method. With S_{ah} and S_{bh}, the amplitude a_h and phase φ_h of the h^{th} order residual system harmonic current can be calculated according to Eqs. (96.18) and (96.19). It should be noted that a_h and φ_h are calculated only once at each control cycle:

$$a_h = \frac{2\sqrt{S_{ah}{}^2 + S_{bh}{}^2}}{N} \tag{96.18}$$

$$\varphi_h = \begin{cases} \arctan(S_{ah} / S_{bh}) & S_{bh} > 0 \\ \arctan(S_{ah} / S_{bh}) + \pi & S_{bh} < 0 \end{cases} \tag{96.19}$$

Therefore, according to Eqs. (96.16) to (96.19), the output voltage correction signal of APF can be calculated.

96.4 Simulation and Analysis

Fig. 96.3 shows the control system block diagram of the parallel hybrid filters, including harmonic control and direct current (DC) voltage control.

The harmonic control section consists of calculation of the proportional control signal and the correction signal. The harmonic current can be obtained using system current and the fundamental current. The proportional control signal is k times of the harmonic current. Output voltage correction signal of APF is calculated using Eqs. (96.16) to (96.19).

The inputs of the DC voltage control section are the grid voltage U_{S} and DC voltage U_{d}. The output active current of APF is only related to the component of APF output voltage which is perpendicular to the grid voltage [3]. By adjusting this component with the difference between the set value and true value of DC voltage, DC voltage can be regulated stably.

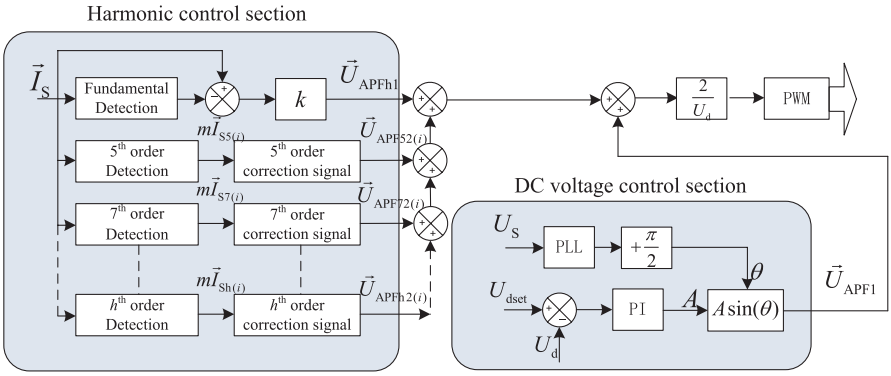


Fig. 96.3 Control system block diagram of shunt HPF

The response time of the sliding window iterative DFT method is about the width of the sliding window, which is 20 ms here. Therefore the control cycle is selected as 20 ms. PSCAD simulation software is used to create a model based on Fig. 96.3. The simulation parameters are as follows: system line root mean squared (RMS) voltage 380 V, system frequency 50 Hz, system inductance $L_S = 0.22$ mH; PF harmonic tunes at 237.5 Hz, $C = 300$uF, $L_{PF} = 1.5$ mH, $R_{PF} = 0.0471\ \Omega$, with a quality factor of 50; APF DC capacitor $C_d = 5000$ uF, DC voltage setting is 150 V, APF output inductance $L_{APF} = 1.41$ mH, resistance $R_{APF} = 0.05\ \Omega$; current sampling frequency and the PWM carrier frequency is 12800 Hz.

With sufficient stability margin, $k = 25$ and $m = 25$ are chosen for simulation. Simulations are conducted with and without background harmonic; the considered background harmonic are 5th, 7th, and 11thwith the value of 3, 2 and 1 % ,respectively.

HPF is put into operation at 0.5 s, and load doubled at 1.5 s. Figure 96.4 shows waves of phase-A system current and load current during 0.5 s both with and without background harmonic, and Fig. 96.5 shows the same waves during 1.5 s.

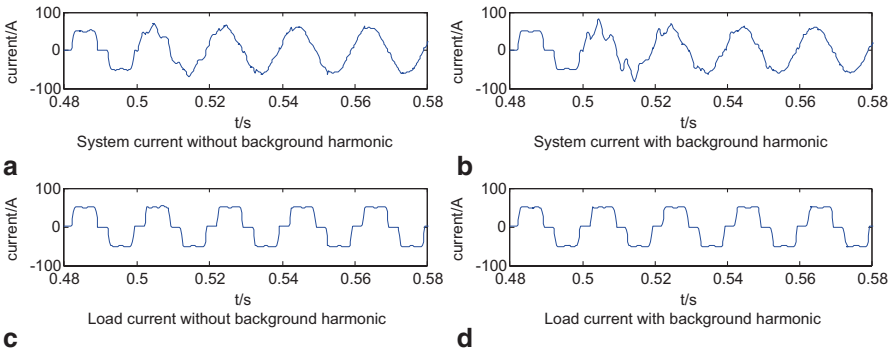


Fig. 96.4 System current and load current during 0.5s

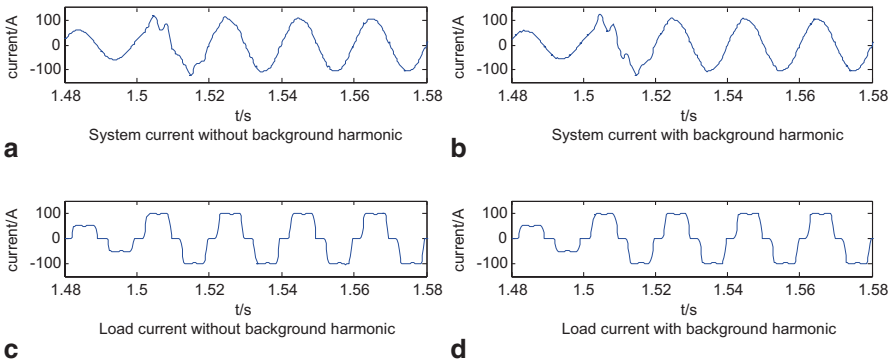


Fig. 96.5 System current and load current during 1.5s

Table 96.1 The main harmonic content of the load current and system current

Order	5th (A)	7th (A)	11th (A)	13th (A)
Load current	15.49	8.17	4.49	3.2
System current (without background harmonic)	0.03	0.061	0.051	0.088
System current (with background harmonic)	0.042	0.083	0.082	0.079

Table 96.1 shows the main harmonic content of the load current and system current both with and without background harmonic.

It can be seen from Fig. 96.4 that the system current is almost a sinusoidal signal one or two cycles after HPF is put into operation. Figure 96.5 shows that the system current recovers to almost a sinusoidal signal one cycle after load doubled. Table 96.1 makes it clear that the filtering rate is high with the proposed method. The results indicate that the proposed method performs well either with or without background harmonic.

96.5 Conclusion

A control strategy based on output voltage correction for the HPF is proposed in this chapter. On the basis of proportional control, a correction sinusoidal signal with certain amplitude and phase is calculated for APF output voltage to eliminate harmonics. A simple but effective method is given to determine the parameters of the control strategy, and the calculation method of the amplitude and phase of the correction signal are also presented. The simulation results show that the proposed method performs well with high filtering rate either under background harmonic or without.

References

1. Park S, Sung J, Nam K. A new parallel hybrid filter configuration minimizing active filter size[C]. 30th IEEE Annual Power Electronics Specialists Conference, IEEE; 1999. pp. 400–5.
2. Detjen D, Jacobs J, De Doncker RW, et al. A new hybrid filter to dampen resonances and compensate harmonic currents in industrial power systems with power factor correction equipment [J]. IEEE Trans Power Electr. 2001;16(6):821–7.
3. Curi BTD, Vendrusculo EA, Antenor PJ, et al. Analysis of a derivative hybrid power filter in distorted voltage grid [C]. 2013 IEEE PES Conference On Innovative Smart Grid Technologies Latin America, IEEE; 2013. pp. 1–5.
4. Asiminoaei L, Wiechowski W, Blaabjerg F, et al. A new control structure for hybrid power filter to reduce the inverter power rating [C]. 32nd IEEE Industrial Electronics Annual Conference, IEEE; 2006. pp. 2712–7.
5. Valdez AA, Escobar G, Martinez-Montejano MF. A model-based controller for a hybrid power filter to compensate harmonic distortion in unbalanced operation [C]. IEEE Power Electronics Specialists Conference, IEEE; 2008. pp. 3861–6.
6. SHI X, Wang Z, Zhang J, et al. Study on the control strategy for single-tuned hybrid active power filters [J]. Electr Mach Contr. 2012;16(4):64–71. (In Chinese).
7. Yu Y, Xu Y, Liu X. Study of harmonic current detection based on sliding-window iterative algorithm of DFT [J]. Power Syst Prot Contr. 2011;39(13):78–82. (In Chinese).

Chapter 97
Multiscale Relevance Vector Machine Fault Prediction Based on Genetic Algorithm Optimization

Yongliang Zhang, Guolin Li and Cuihua Lu

Abstract As the present fault prediction methods are not accurate enough, a multiscale relevance vector machine (MSRVM) based on genetic algorithm (GA) optimization is proposed. The kernel scales' number and kernel parameters are optimized by GA to improve the performance of the MSRVM. Its feasibility and advantages are proved by the fault prediction of a Buck converter circuit

Keywords Fault prediction · MSRVM · Optimization · GA

97.1 Introduction

The fault prediction is a more advanced maintenance type than the fault prognosis [1], and it is the core content of prognostics and health management (PHM) of the equipment [2]. It can estimate the remaining useful life of the units or system and predict their future health conditions with the comprehensive utilization of all varieties of data, such as the parameters monitored, use state, current environment and working condition, early experiment data, and historical experience while assisted by all varieties of reasoning technologies, such as mathematics–physics models and artificial intelligence [3]. The fault prediction methods can be classified as the model-driven based method, the data-driven based method, and the reliability and statistics based method [4], from among which the data-driven based method has a broad application perspective because it can predict fault with the test data or the data from the sensors and does not need to build the model of the units or system.

Y. Zhang (✉) · G. Li · C. Lu
Naval Aeronautics and Astronautics University, 264001 Yantai, China
e-mail: yyyh1@sina.com

G. Li
e-mail: yyyh1@sina.cn

C. Lu
e-mail: yyyh2@sina.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_97

At present, the main data-driven based methods are time series [5], neural network [6], support vector machine (SVM) [7], and so on. The time series is not accurate enough for the nonlinearity prediction. The neural network is likely to fall into local optimum and slow in convergence speed. SVM also has many shortcomings; for example, its model is not sparse enough, output is nonprobabilistic, parameters are too complex to set, and kernel functions have to satisfy Mercer's condition. The relevance vector machine (RVM), originally introduced by Tipping [8], is a Bayesian learning model, which provides the state-of-the-art results both in terms of accuracy and sparsity, overcomes all of the shortcomings of the SVM above, and has led to many successful applications for fault prediction within a certain scope [9, 10]. But the limitation of RVM with a single kernel has blocked the improvement of the prediction accuracy [11], which can be solved by the method of MSRVM.

Therefore, a multiscale relevance vector machine (MSRVM) for fault prediction is proposed in this chapter. Its kernel scales' number and kernel parameters are optimized by genetic algorithm (GA) to improve the prediction accuracy of the RVM model, then the feasibility and advantages are proved by the fault prediction of a Buck converter circuit.

97.2 MSRVM Regression

97.2.1 Model Specification

Given is a data set of input-target training pairs $\{\boldsymbol{x}_n, t_n\}_{n=1}^{N}$, where the input vector $\boldsymbol{x}_n \in \mathbf{R}^d$ and the targets $t_n \in \mathbf{R}$. Assume that the targets are independent, and samples from the model with additive noise

$$t_n = y(\boldsymbol{x}_n, \boldsymbol{w}) + \varepsilon_n, \tag{97.1}$$

where ε_n are independent samples from some noise process which is further assumed to be mean-zero Gaussian with variance σ^2; then RVM makes predictions based on the function

$$y(\boldsymbol{x}, \boldsymbol{w}) = \sum_{i=1}^{N} w_i K(\boldsymbol{x}, \boldsymbol{x_i}) + w_0, \tag{97.2}$$

where $K(\boldsymbol{x}, \boldsymbol{x}_i)$ is a kernel function, w_i is the weight factor, and N is the number of samples. Then the likelihood of the complete data set can be written as

$$p(\boldsymbol{t} \mid \boldsymbol{w}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}\left\|\boldsymbol{t} - \boldsymbol{\Phi w}\right\|^2\right\}, \tag{97.3}$$

where $\boldsymbol{t} = [t_1 \cdots t_N]^{\mathrm{T}}$, $\boldsymbol{w} = [w_0 \cdots w_N]^{\mathrm{T}}$, $\boldsymbol{\Phi}$ is the $N \times (N+1)$ "design" matrix with $\boldsymbol{\Phi} = [\boldsymbol{\phi}(\boldsymbol{x}_1), \boldsymbol{\phi}(\boldsymbol{x}_2), \cdots, \boldsymbol{\phi}(\boldsymbol{x}_N)]^{\mathrm{T}}$, where in $\boldsymbol{\phi}(\boldsymbol{x}_n) = [1\ K(x_n, x_1)\ \cdots\ K(x_n, x_N)]^{\mathrm{T}}$.

With many parameters in the model as training examples, the maximum-likelihood estimation of $\boldsymbol{w}$ and σ^2 will lead to severe overfitting. To avoid this, a sparse Bayesian method is adopted to define a mean-zero Gaussian prior distribution over $\boldsymbol{w}$:

$$p(\boldsymbol{w} \mid \boldsymbol{\alpha}) = \prod_{i=0}^{N} N(w_i \mid 0, \alpha_i^{-1}), \tag{97.4}$$

with α as a vector of $N + 1$ hyperparameters.

97.2.2 Bayesian Inference

From Bayes' rule, the posterior over all unknowns given the data

$$p(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2 \mid \boldsymbol{t}) = \frac{p(\boldsymbol{t} \mid \boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2)}{p(\boldsymbol{t})}. \tag{97.5}$$

Then, given a new test point, $\boldsymbol{x}_*$, predictions are made for the corresponding target t_*, in terms of the predictive distribution:

$$p(\boldsymbol{t}_* \mid \boldsymbol{t}) = \int p(\boldsymbol{t}_* \mid \boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2 \mid \boldsymbol{t}) \mathrm{d}\boldsymbol{w} \mathrm{d}\boldsymbol{\alpha} \mathrm{d}\sigma^2. \tag{97.6}$$

$p(w, \alpha, \sigma^2 \mid t)$ in (97.5) cannot be computed directly, considering

$$p(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2 \mid \boldsymbol{t}) = p(\boldsymbol{w} \mid \boldsymbol{t}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}, \sigma^2 \mid \boldsymbol{t}). \tag{97.7}$$

Then the posterior distribution over the weights is given by:

$$p(w \mid t, \alpha, \sigma^2) = \frac{p(t \mid w, \sigma^2) p(w \mid \alpha)}{p(t \mid \alpha, \sigma^2)} = (2\pi)^{-(N+1)/2} \left|\Sigma\right|^{-1/2} \exp\left\{-\tfrac{1}{2}(w-\mu)^T \Sigma^{-1} (w-\mu)\right\}, \tag{97.8}$$

where the posterior covariance is $\boldsymbol{\Sigma} = (\sigma^{-2}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi} + \boldsymbol{A})^{-1}$ and the mean is $\boldsymbol{\mu} = \sigma^{-2}\boldsymbol{\Sigma}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{t}$ with $\boldsymbol{A} = \mathrm{diag}(\alpha_0, \alpha_1, \cdots, \alpha_N)$.

The hyperparameters posterior $p(\boldsymbol{\alpha}, \sigma^2 \mid \boldsymbol{t})$ in (97.7) are approximated by the peak of a delta function, i.e., at its most probable values $\boldsymbol{\alpha}_{\mathrm{MP}}$ and σ^2_{MP}. For predictive purposes, rather than requiring $p(\boldsymbol{\alpha}, \sigma^2 \mid \boldsymbol{t}) \approx \delta(\boldsymbol{\alpha}_{\mathrm{MP}}, \sigma^2_{\mathrm{MP}})$, only to desire

$$\int p(t_* \mid \alpha, \sigma^2) \delta(\alpha_{MP}, \sigma^2_{MP}) d\alpha \, d\sigma^2 \approx \int p(t_* \mid \alpha, \sigma^2) p(\alpha, \sigma^2 \mid t) d\alpha \, d\sigma^2 \tag{97.9}$$

to be a good approximation.

The relevance vector "learning" thus becomes the search for the mode. As for the case of uniform hyperpriors, the only need is to maximize the term $p(\boldsymbol{t} \mid \boldsymbol{\alpha}, \sigma^2)$,

$$\begin{aligned} p(t\left|\alpha,\sigma^2\right.) &= \int p(t\left|w,\sigma^2\right.)p(w\left|\alpha\right.)dw \\ &= (2\pi)^{-N/2}\left|\sigma^2 I + \Phi A^{-1}\Phi^T\right|^{-1/2} \exp\left\{-\frac{1}{2}t^T(\sigma^2 I + \Phi A^{-1}\Phi^T)^{-1}t\right\}. \end{aligned} \tag{97.10}$$

97.2.3 Optimizing the Hyperparameters and Making Predictions

Values of α and σ^2 which maximize (97.10) cannot be obtained in a closed form, and here we summarize the formula for their iterative re-estimation, $\alpha_i^{\text{new}} = \dfrac{\gamma_i}{\mu_i^2}$ and $(\sigma^2)^{new} = \dfrac{\|\boldsymbol{t} - \boldsymbol{\Phi\mu}\|}{N - \Sigma_i \gamma_i}$,

where μ_i is the i-th posterior mean weight and $\gamma_i \equiv 1 - \alpha_i \Sigma_{ii}$ with Σ_{ii} as the i-th diagonal element of the posterior weight covariance.

The estimation of the weight factors in RVM is given by the mean of the posterior distribution, which is also the maximum posterior distribution of the weight factors. The posterior distribution can reflect the uncertainty of the optimal value of the weight factors and of the predictions of the model.

Compute the predictive distribution for a new datum $\boldsymbol{x}_*$:

$$p(t_* \mid \boldsymbol{t}, \boldsymbol{\alpha}_{\text{MP}}, \sigma^2_{\text{MP}}) = \int p(t_* \mid \boldsymbol{w}, \sigma^2_{\text{MP}}) p(\boldsymbol{w} \mid \boldsymbol{t}, \boldsymbol{\alpha}_{\text{MP}}, \sigma^2_{\text{MP}}) \mathrm{d}\boldsymbol{w}. \tag{97.11}$$

Both terms in the integrand are Gaussian, which is readily computed as below:

$$p(t_* \mid \boldsymbol{t}, \boldsymbol{\alpha}_{\text{MP}}, \sigma^2_{\text{MP}}) = N(t_* \mid y_*, \sigma_*^2), \tag{97.12}$$

with $y_* = \boldsymbol{\mu}^{\text{T}} \phi(\boldsymbol{x}_*)$ and $\sigma_*^2 = \sigma^2_{\text{MP}} + \phi(\boldsymbol{x}_*)^{\text{T}} \boldsymbol{\Sigma} \phi(\boldsymbol{x}_*)$; so the predictive mean is $y_* = (\boldsymbol{x}_*; \boldsymbol{\mu})$.

97.2.4 MSRVM Regression Modeling

The Gaussian kernel function is a local kernel function with strong learning ability but weak generalization performance, which can be changed obviously by selecting different kernel parameters; in other words, which makes the Gaussian kernel function to be a typical one that can be multiscale. To take it as the kernel function of RVM, the multiscale Gaussian kernel function built in this chapter is

$$K(x,z) = \sum_{i=1}^{n} \lambda_i \cdot \exp\left(-\frac{\|x - z\|^2}{2\gamma_i^2}\right), \tag{97.13}$$

where λ_i and γ_i are the kernel parameters to be determined, and $\lambda_i \in [0,1]$.

Take the above multiscale Gaussian kernel function into RVM modeling.

97.3 Kernel Scales' Number and Kernel Parameters Optimization Based on GA

There is no clear standard for the selection of the scales' number. An ideal prediction cannot be obtained if the number is too small while the computational complexity will increase largely when the number is too high. The values of the kernel parameters will affect the performance of the combination kernel function, and then affect the performance of MSRVM; therefore, GA is adopted to optimize the kernel scales' number and the kernel parameters in this chapter.

97.3.1 Design of GA

The segmented binary encoding is adopted here. Suppose there are n scales at most in MSRVM; then each chromosome code can be equally divided into $2n + 1$ segments including $2n$ segments of the kernel parameter code and one segment of scales' number code where a segment of the kernel parameters code is composed of a weight factor code and its corresponding kernel width code. The length of the weight factor code is $m1$, the length of the kernel width code is $m2$, and the length of scales' number code is $m3$, which is the shortest encoding length for n to transform into binary. Then the segmented binary decoding is adopted when the encodings are decoded.

Make use of the mean relative error (MRE) to measure the prediction accuracy of MSRVM; then the MRE can be expressed as

$$\delta = \frac{1}{l}\left(\sum_{i=1}^{l}\left(\left|\hat{y}_i - y_i\right| / y_i\right) \times 100\%\right), \tag{97.14}$$

where $\hat{y}$ is the prediction value, y is the real value, and l is the data length.

The smaller MRE is, the higher the prediction accuracy of MSRVM will be. This is a problem to minimize the target function. Then the fitness function can be expressed as

$$F(x) = 1/\delta. \tag{97.15}$$

The selection operator is a proportional selection, the crossover operator is two points crossover, and the mutation operator is a uniform mutation.

97.4 Analysis of the Prediction Results

The performance of the method in this chapter (marked as $MSRVM_a$) will be tested with the data [6], and its prediction results will be compared with the ones obtained by the least squares support vector machine (LSSVM) [6], 3-scale RVM ($MSRVM_3$) including $MSRVM_{3r}$ whose kernel parameters are set randomly and $MSRVM_{3a}$ whose kernel parameters are optimized by GA, 6-scale RVM ($MSRVM_6$). As to 10 time points, the ripple voltage values are 0.7680, 0.7820, 0.7930, 0.8090, 0.8250, 0.8420, 0.8580, 0.8890, 0.9050, and 0.9280. Take the first six values of the ripple voltage as the first training sample, with which to predict the 7th value. Then make up the second training sample with the values on the 2–7th time point, with which to predict the 8th value, and the same as to the values on other time points left to be predicted. Taking account of the efficiency and performance of MSRVM, suppose $n = 6$, $\gamma_i \in [0,10]$, $m1 = 4$, $m2 = 6$, $m3 = 3$, population number M = 80, generation T = 100, crossover probability $p_c = 0.7$, mutation probability $p_m = 0.01$. The program is running under MATLAB 8.0. The evolution progress of GA is shown in Fig. 97.1 and the prediction results of the ripple voltage are shown in Table 97.1.

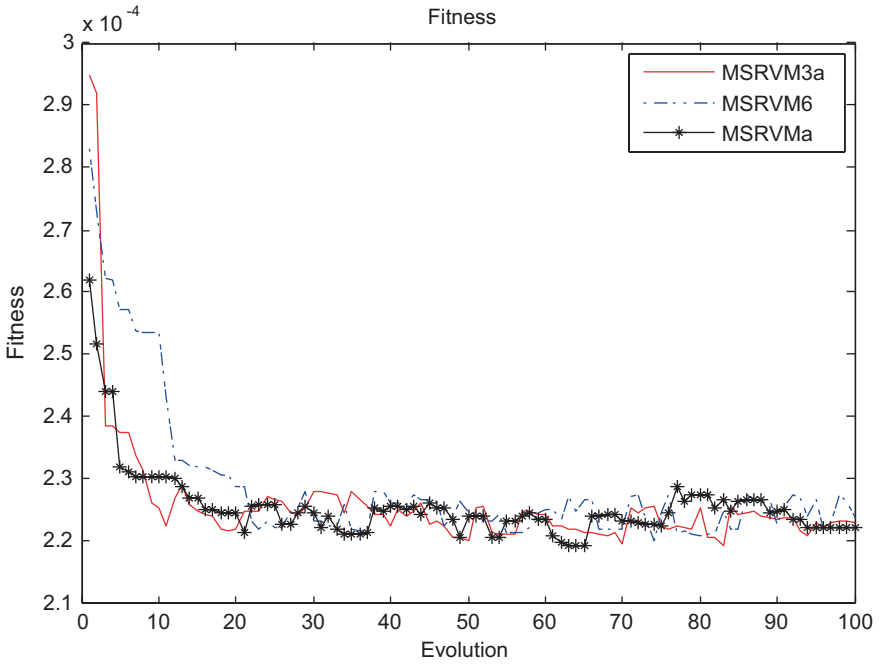


Fig. 97.1 Evolution progress of GA

Table 97.1 Prediction results of ripple voltage

Time point/h	Real value/V	Prediction value/V				
		LSSVM	$MSRVM_{3r}$	$MSRVM_{3a}$	$MSRVM_6$	$MSRVM_a$
7	0.8580	0.8456	0.8916	0.8579	0.8579	0.8577
8	0.8890	0.8724	0.8943	0.8746	0.8747	0.8744
9	0.9050	0.9003	0.8963	0.9095	0.9095	0.9092
10	0.9280	0.9234	0.8971	0.9291	0.9291	0.9288
Relative error (%)		0.3963	3.9161	0.0119	0.0062	0.0400
		1.8673	0.5962	1.6178	1.6119	1.6459
		0.5193	0.9613	0.4960	0.4995	0.4655
		0.4957	3.3297	0.1132	0.1161	0.0829
Mean relative error (%)		0.8196	2.2008	0.5597	0.5584	0.5586
Max relative error (%)		1.8673	3.9161	1.6178	1.6119	1.6459
Min relative error (%)		0.3963	0.5962	0.0119	0.0062	0.0400

After optimization by GA, the scales' number of $MSRVM_a$ is 4, and the kernel parameters of $MSRVM_{3a}$, $MSRVM_{6,}$ and $MSRVM_a$ are $\lambda_{3a} = [0.4000, 0.4667, 0.2000]$, $\gamma_a = [7.7778, 9.5238, 9.0476]$, $\lambda_a = [0.4000, 0.1333, 0.0667, 0.2000, 1.0000, 0.4667]$, $\gamma_a = [9.3651, 5.8730, 8.0952, 9.6852, 9.5238, 9.3651]$, $\lambda_a = [0.3333, 0.0667, 0.4667, 0.2000]$, and $\gamma_a = [9.3651, 6.1905, 8.7302, 7.4603]$, respectively

1. The prediction results in Table 97.1 show that, the values of ripple voltage in the future four time points will change within the limits of threshold; thus the circuit is predicted to be normal during this period, which is in conformity with the fact.
2. Comparing the results of $MSRVM_{3a}$ and $MSRVM_{3r}$ shows that there is one order of magnitude at least in the prediction results with the results of the former being more accurate than the ones of the latter. It indicates that the prediction performance of the model can be enhanced obviously through optimization of the kernel parameters; thus, it is of great necessity to optimize the kernel parameters.
3. Comparing the results of $MSRVM_a$, $MSRVM_{3a}$, and $MSRVM_6$ shows that $MSRVM_a$ is close to $MSRVM_6$ but higher than $MSRVM_{3a}$ in the prediction accuracy, while $MSRVM_a$ is close to $MSRVM_{3a}$ but less than $MSRVM_6$ in the kernel scales' number; therefore, the overall performance of $MSRVM_a$ is the best of all. It indicates that the most proper number range of the kernel scales for $MSRVM_a$ to predict the voltage values of the Buck converter circuit is 4 ~ 6 and also proves the necessity of the kernel parameters optimization.
4. Comparing the results of $MSRVM_a$ and LSSVM shows that $MSRVM_a$ is more accurate than LSSVM in terms of prediction accuracy, especially that the former's smallest relative error is very close to 0, which LSSVM cannot achieve; therefore, the feature parameters' change tendency of the Buck converter circuit can be predicted more accurately by $MSRVM_a$ than before.

97.5 Conclusion

In this chapter, a MSRVM based on GA optimization for the fault prediction is proposed; then the feasibility and advantages of MSRVM are proved by the fault prediction of the Buck converter circuit. There will be more application ranges for MSRVM and more real application values therein in the future because of its advantages in terms of prediction accuracy and efficiency in terms of training and prediction.

The kernel scales' number and kernel parameters are optimized by GA, and the kernel functions are Gaussian kernel functions only; besides, more attention should be paid to the MSRVM with different kernel functions based on other methods of optimization.

References

1. Liu CS, Hu SS. Intelligent nonlinear fault diagnosis based on state estimator [J]. Contr Decis. 2005;20(5):557–61.
2. Keith MJ, Raymond RB. Diagnostics to prognostics: a product availability technology evolution [C]. The 53rd Annual Reliability and Maintainability Symposium (RAMS 2007), Orlando, FL, USA; 2007. pp. 113–8.
3. Hess A, Fila L. Prognostics, from the need to reality from the fleet users and PHM system designer/developers perspectives [C]. Aerospace Conference Proceedings, IEEE; 2002. pp. 2791–7.
4. Hess AC, Frith GP. Challenges, issues, and lessons learned chasing the "Big P". Real predictive prognostics. Part1 [C]. IEEE Aerospace Conference Proceedings, Big Sky, Montana, USA; 2005. pp. 3610–9.
5. Lv K, Qiu J, Liu G. Research on life prognosis method for electronics based on dynamic damage and optimization ar model [J]. Acta Armamentarii. 2009;30(1):91–5. (In Chinese).
6. Jiang Y, Wang Y, Cui J, et al. Research on fault prediction method of power electronic circuits based on least squares support vector machine [J]. Electr Mach Contr. 2011;15(8):64–8. (In Chinese).
7. Ma D, Fan G, Zhang J. Relevance vector machine and its applications in fault diagnosis and prognosis [J]. J Nav Aeronaut Astronaut Univ. 2013;28(2):154–60. (In Chinese).
8. Tipping ME. Sparse Bayesian learning and the relevance vector machine [J]. J Mach Learn Res. 2001;1:211–44.
9. Saha B, Goebel K, Poll S, et al. Prognostics methods for battery health monitoring using a Bayesian framework [J]. IEEE Trans Instrum Meas. 2009;58(2):291–6.
10. Caesarendra W, Widodo A, Yang B. Application of relevance vector machine and logistic regression for machine degradation assessment [J]. Mech Syst Signal Process. 2010;60(1):1161–71.
11. Wang H, Sun F, et al. On multiple kernel learning methods [J]. Acta Autom Sin. 2010;85:1037–50. (In Chinese).

Chapter 98
Automotive Light Control System Development Based on Kinetis KEA

Feng Luo and Yanwei Zhai

Abstract In this chapter, an automotive light control system based on the Local Interconnect Network (LIN) network is developed. It consists of one master node and four slave nodes with a 32-bit microcontroller (MCU) Kinetis KEA64 adopted as the core microprocessor; besides, the hardware design mainly includes the discussion of the characteristics and advantages of Kinetis KEA64 as well as the hardware architecture of the system, and the software design introduces the software architecture and implements the LIN driver and software flow design. Finally, the system can realize basic light control functions. At the moment of testing the LIN network with the Emulin hardware tool, the system proves to be reliable, by which both the communication's real timing and the dependability can be guaranteed.

Keywords Automotive light control · LIN · Kinetis KEA · Emulin

98.1 Introduction

Intelligent automotive light control systems such as the Adaptive Front-Lighting System (AFS) can improve the performance of automotive lighting system and increase driving safety [1]. It becomes more and more important with the development of automotive electronics. The Local Interconnect Network (LIN) is a widely used network in vehicle-body control such as in the automotive light control system; the LIN bus is a single master with a multiple-slave network based on common universal asynchronous receiver/transmitter (UART)/serial communication interface (SCI) hardware; and it can be single-wire implemented with a speed of up to 20 kbps. It can be a more cost-effective bus and a sub-bus of controller area network (CAN). LIN also implements the transport layer and supports diagnosis [2]. Kinetis KEA64 (Kinetis E-series Automotive) is a newly designed 32-bit automotive grade microcontroller (MCU) by Freescale Inc. based on ARM architecture. In automotive body application, the Kinetis Auto family is a great option for the entry level

F. Luo (✉) · Y. Zhai
Clean Energy Automotive Engineering Center, School of Automotive Studies, Tongji University, 201804 Shanghai, China
e-mail: luo_feng@tongji.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_98

body controller or gateway module [3]. This chapter works with an automotive light control system development, which uses Kinetis KEA as the core processor and LIN as the communication network. Due to high performance of the ARM architecture MCU and the low cost of LIN, the system is highly significant for further developing an advanced vehicle-body electronics system.

98.2 System Design

The system architecture is depicted in Fig. 98.1. The automotive light control system uses the distributed structure [3]. The master node of the LIN bus is equivalent to a central control unit, and four slave nodes control the corresponding automotive light module [4]. As depicted in Fig. 98.1, each light module consists of several lamps, which are positioned according to the layout of the car.

This automotive light control system is to control and diagnose the steering lamp, braking lamp, reversing lamp, and headlamp. Users can control the light modules by pushing switches on a switch panel or clicking corresponding buttons on the PC visual interface. Then the master node sends related message onto the LIN bus. The slave nodes receive message from the LIN bus and control the corresponding lamps through a power driver module.

As depicted in Fig. 98.1, the master node and four slave nodes are connected through a single wire of LIN. The power driver module is used to output high voltage to drive the lamps; each power driver has four output channels at most. All lamps are the same as those used on a GEELY car, and they are 20 and 55 W.

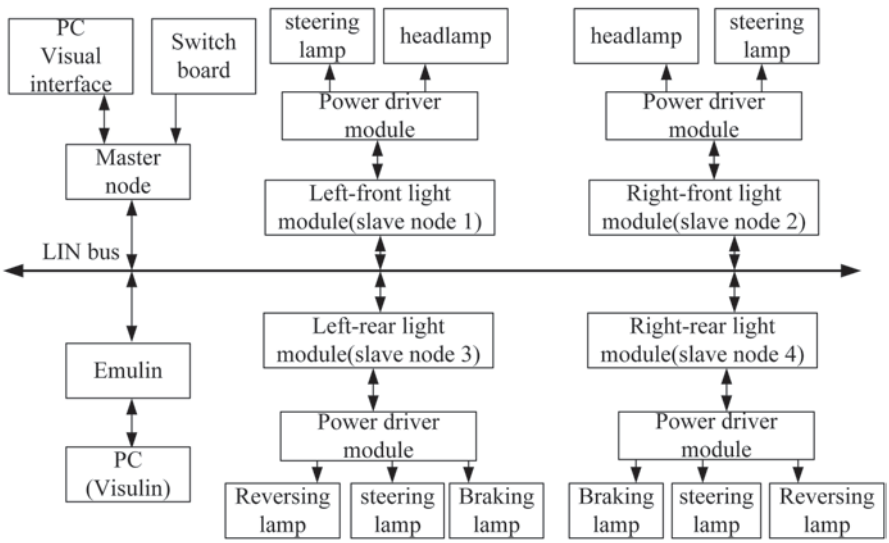


Fig. 98.1 System architecture

The Emulin hardware testing tool and Visulin software are developed by Ihr GmbH, Germany, as professional testing tools. Emulin can be configured as a master node of the LIN bus to control the whole network or configured as a slave node to monitor the LIN bus. Visulin provides users an interface to configure the parameters of Emulin. Both Emulin and Visulin are employed into the system for testing and validating the network functions and showing a message on the LIN bus.

98.2.1 Hardware Design

98.2.1.1 Kinetis KEA

Kinetis KEA is a newly designed automotive grade MCU based on 32-bit ARM Cortex-M0+ MCUs aiming for the automotive markets. It is optimized for cost-sensitive applications offering a low-pin count option with very low power consumption with a 2.7–5.5 V supply. Kinetis Auto series devices are well suitable for a wide range of applications ranging from body applications and safety to generic sensor nodes. All the members of the Kinetis Auto family share similar peripherals and offer several pin count and memory options allowing developers to migrate easily to MCUs that take advantage of more memory or peripheral integration [5]. Also a lot of third-party software supports Kinetis Auto series MCUs, including CodeWarrior, KEIL, IAR, Processor Expert, and MQX. Kinetis KEA64 provides users many modules including an up to 64-K flash, various kinds of timers, AD conversion module, GPIOs, and keyboard interrupt. KEA64 also integrates many communication modules such as serial peripheral interface (SPI), I2C, UART, etc.

ARM is a well-known architecture for microprocessors, whose power consumption is very low but performance is high because of its instruction set. Also the low pin count and low price with many communication source modules makes KEA64 suitable for this automotive light control system. Its UART module supports LIN break detection, and this characteristic is used in LIN driver implementation.

98.2.1.2 Hardware Architecture

The hardware architecture is described in Fig. 98.2. Similar to the system architecture, the hardware architecture consists of one master node and four slave nodes each for one light module.

Each node has a MC33662 LIN transceiver, whose basic function is to convert the transistor–transistor logic (TTL) voltage of the MCU from and to the voltage of the LIN bus. MC33662 offers a normal baud rate of 20kbps. It is compatible with LIN protocol specification 1.3/2.0/2.1 and SAE J2602-2. MC33662 also has safe behavior in case of a LIN bus short-to-ground or a LIN bus leakage during low power mode. Digital inputs are 5.0 and 3.3 V compatible without any external required components.

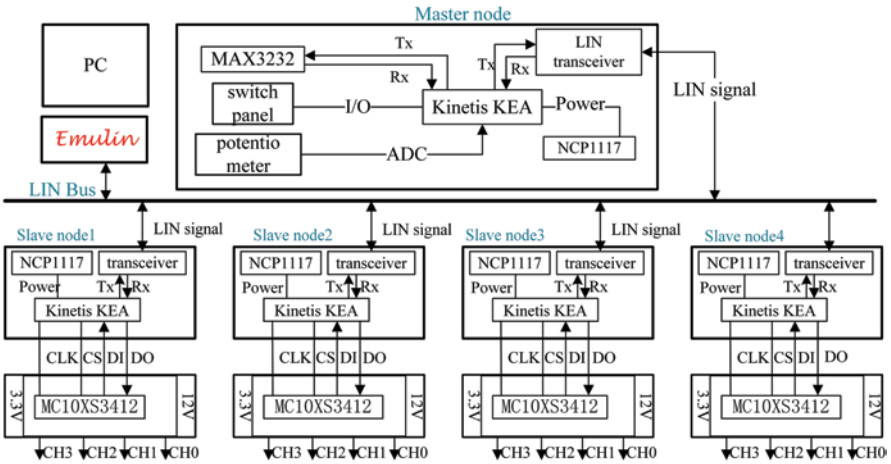


Fig. 98.2 Hardware architecture

Each slave node works with a power driver module, whose main chip is MC10XS3412. MC10XS3412 is designed for low-voltage automotive lighting applications. Its four low $R_{DS(ON)}$ MOSFETs (dual 10 mΩ/dual 20 mΩ) can control four separate 55/28 W bulbs, and/or Xenon modules, and/or LEDs. Programming, control, and diagnostics can be accomplished by using a 16-bit SPI interface. Additionally, each output has its own parallel input or SPI control for pulse width modulation (PWM) control if desired. MC10XS3412 has many protections including smart over-current shutdown, severe short-circuit, over-temperature protection, etc. Output OFF or ON open-load detection compliant to bulbs or LEDs and short to battery detection, the analog current feedback with selectable ratio, and the board temperature feedback can be all realized by MC10XS3412. In this automotive light control system, the diagnosis of headlamps is just based on the Output OFF or ON open-load detection.

98.2.2 Software Design

Figure 98.3 shows the software architecture, which has a layer structure (refer to AUTOSAR) [6]. The topmost layer is an application component layer, including some application control module such as the analog to digital converter (ADC) control module and drive board control module. The hardware abstraction layer is on top of the processor hardware; it consists of low-level drive codes including input/output (I/O) abstraction, ADC abstraction, SPI abstraction, and UART/SCI abstraction. Communication protocols such as LIN driver and SPI driver are based on the hardware abstraction, which may simplify the software design and only supplies users with necessary application program interfaces (APIs).

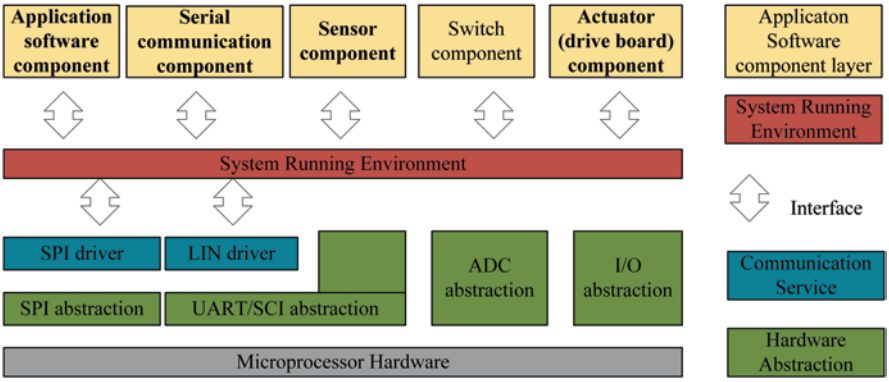


Fig. 98.3 Software architecture

Figure 98.5 shows that the LIN Task Behavior Model LIN driver hides the implementation of the LIN network and only provides users with necessary APIs. This simplifies and accelerates the software development. The LIN driver consists of three parts including LIN API, LIN configuration (CFG), and LIN protocol. LIN API provides users with the application interface functions; LIN CFG is used for setting the LIN network parameters; LIN protocol is an implementation of the LIN protocol with many functions and a series of state machines [7]. The structure of the LIN driver is shown in Fig. 98.4.

A key property of the LIN protocol is the use of schedule tables which makes it possible to assure that the bus will never be overloaded. The master node is responsible for the task schedule table as defined in the LIN configuration in a specific period of time. The master task in the master node is responsible for sending the

Fig. 98.4 LIN driver structure

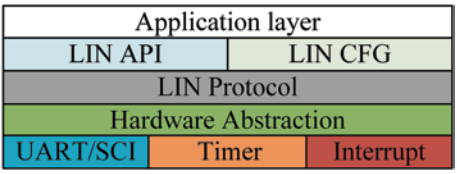
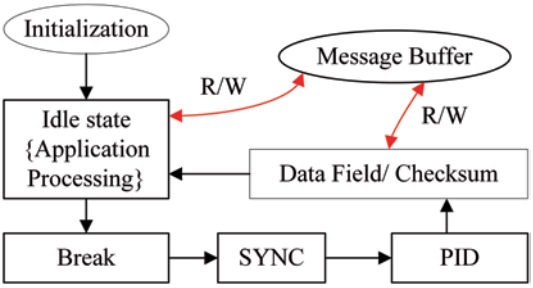


Fig. 98.5 LIN Task Behavior Model



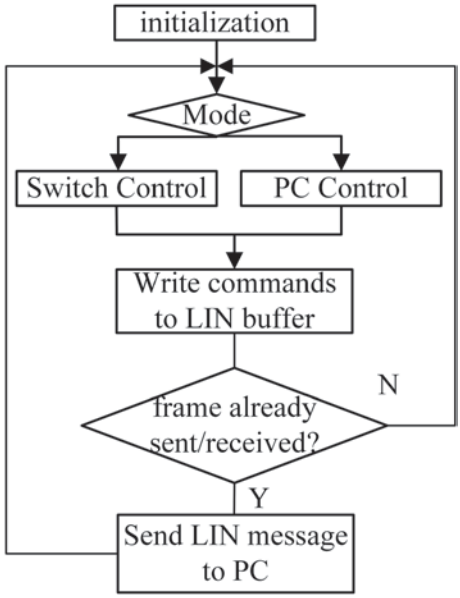


Fig. 98.6 Master Node Software Flow

break field, sync field, and PID (Parity ID, ID: Identifier) related to each message. The slave tasks in the master node or slave nodes respond to the header field by receiving or sending the data field according to each PID. Both the master node and the slave nodes process the application layer code during the idle state. The LIN task behavior model is described in Fig. 98.5.

In the idle state, the master node scans whether it is in switch control mode or in PC control mode. Then the master node reads the commands and writes them to the LIN buffer. After the message has been sent or received, it should also be sent to the PC through UART/SCI. The master node's application software flow is described in Fig. 98.6.

Similar to the master node, the slave node first checks the LIN signal status. When the command signals have been received, the slave node reads signals from the LIN buffer and writes commands to the drive board through SPI. Then the slave node reads back the status of MC10XS3412 and writes them to the LIN buffer. The application software flow of the slave nodes is depicted in Fig. 98.7.

Five unconditional frames are defined in this system with one control frame containing all controlling signals and four status frames used to return the corresponding light module's status. The control frame is shown in Table 98.1. It has an ID of 0×15; the status frames have IDs of 0×20, 0×21, 0×30, and 0×31. Taking the status frame 1, for example, it has an ID of 0×20. The status frame 1 is shown in Table 98.2.

The control frame has a length of 6 Bytes, of which each Byte controls one automotive light function. The two least significant bits are used to control: 0b11 denotes turning related lamp on and 0b00 indicates turning related lamp off. In the sixth Byte, the bits 2–4 refer to the brightness adjustment of headlamps.

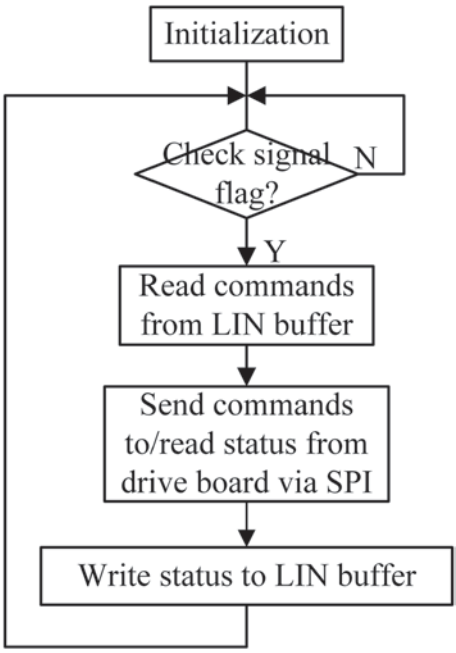


Fig. 98.7 Slave Node Software Flow

Status frame 1 returns the lamps' status of the left-front light module. It has 2 Bytes; the first Byte includes the left-steering lamp status, double flash lamp status, and response error of status frame 1; the second Byte has a left-front headlamp status bit and a headlamp open-load bit indicating whether the headlamp is in open-load status or not.

Table 98.1 Control frame

Bits in Byte		7	6	5	4	3	2	1	0
Control frame (ID = 0 × 15)	Byte 1	Reserved						Left (00:off/11:on)	
	Byte 2	Reserved						Right (00:off/11:on)	
	Byte 3	Reserved						Reversing(00:off/11:on)	
	Byte 4	Reserved						Flash(00:off/11:on)	
	Byte 5	Reserved						Braking(00:off/11:on)	
	Byte 6	Reserved		Headlamp light level				Headlamp(00:off/11:on)	

Table 98.2 Status frame 1

Bits in Octet		7 ~ 3	2	1	0
Status frame1	Byte 1	Reserved	Response error	Flash status	Left status
	Byte 2	Reserved		Headlamp open load	Headlamp status

Fig. 98.8 System Testing Platform

98.3 System Testing

The system testing platform is shown in Fig. 98.8. The validation for the basic light control functions includes the on or off for each lamp and the diagnosis of headlamps. LIN network testing is done with Emulin and Visulin, and LIN physical layer conformance test is done with a scope [8]. The testing results show that the communication between LIN nodes is reliable and real time; meanwhile, the bit time, Byte-space, and frame-space meet the requirements in LIN specification 2.2.

Testing Example 1 Figure 98.9 shows the message displayed in Visulin when the left front headlamp is open load and the right front headlamp is on. According to communication protocol defined in Tables 98.1 and 98.2, the control frame listed in the red box has an ID of 0×15 and includes 6 Bytes. The last Byte 0×07 (0b00000111) indicates that headlamps are being turned on, and the light level is level 2.

The two status frames listed in the green box are the status frames 1and 2, respectively, related to the status of left-front and right-front light modules. Their IDs are 0×20 and 0×21. Both have a data field length of 2 Bytes. The last Byte is designed as an indication for the headlamp. As the left front headlamp is open load, this Byte differs from the right front headlamp. That the left headlamp is not on is shown by 0×02 (0b00000010), and it is in the open-load state.

Testing Example 2 Figure 98.10 shows that five frames are on the LIN network with one control frame subsequently followed by four status frames. Each frame consists of a LIN break field, synchronous field, PID, data field, and checksum.

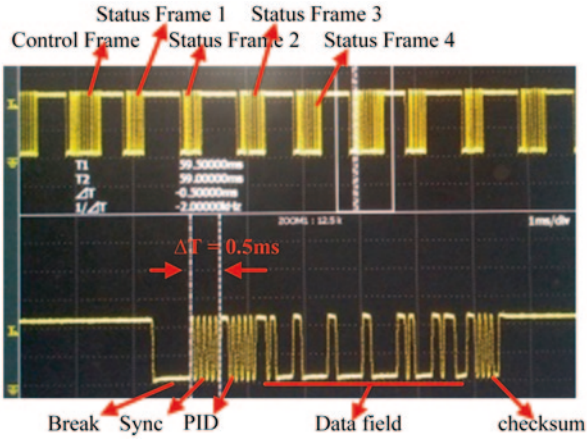
When the UART/SCI baud rate is 19,200 kbps, the bit time is 1 divided by 19,200 about 52 us; thus the time of transmitting one Byte is about 0.52 ms, which agrees well with the test result. The longest frame is the control frame which has about 103 bits. According to LIN specification 2.2A, the minimum time of a frame is 103 multiply by 1.4 multiply by 52 us about 7.5 ms. In this chapter the period of the LIN schedule is 10 ms, which satisfies the requirement.

FrameList1

Timestamp-	ID-	Frame----------	ERR	LE	DATA----------	CHK-
00259.5180	021	RightFrontLampState_Fr	Ok.	02	00 01	[9D]
00259.5292	030	LeftBehindLampState_Fr	Ok.	04	00 00 00 00	[0F]
00259.5392	031	RightBehindLampState_F	Ok.	04	00 00 00 00	[4E]
00259.5504	015	LampControl_Frame1	Ok.	06	00 00 00 00 00 07	[A3]
00259.5580	020	LeftFrontLampState_Fra	Ok.	02	00 02	[DD]
00259.5680	021	RightFrontLampState_Fr	Ok.	02	00 01	[9D]
00259.5792	030	LeftBehindLampState_Fr	Ok.	04	00 00 00 00	[0F]
00259.5893	031	RightBehindLampState_F	Ok.	04	00 00 00 00	[4E]
00259.6005	015	LampControl_Frame1	Ok.	06	00 00 00 00 00 07	[A3]
00259.6080	020	LeftFrontLampState_Fra	Ok.	02	00 02	[DD]
00259.6180	021	RightFrontLampState_Fr	Ok.	02	00 01	[9D]
00259.6292	030	LeftBehindLampState_Fr	Ok.	04	00 00 00 00	[0F]
00259.6392	031	RightBehindLampState_F	Ok.	04	00 00 00 00	[4E]
00259.6505	015	LampControl_Frame1	Ok.	06	00 00 00 00 00 07	[A3]
00259.6580	020	LeftFrontLampState_Fra	Ok.	02	00 02	[DD]
00259.6680	021	RightFrontLampState_Fr	Ok.	02	00 01	[9D]
00259.6792	030	LeftBehindLampState_Fr	Ok.	04	00 00 00 00	[0F]
00259.6893	031	RightBehindLampState_F	Ok.	04	00 00 00 00	[4E]
00259.7005	015	LampControl_Frame1	Ok.	06	00 00 00 00 00 07	[A3]
00259.7080	020	LeftFrontLampState_Fra	Ok.	02	00 02	[DD]
00259.7180	021	RightFrontLampState_Fr	Ok.	02	00 01	[9D]
00259.7292	030	LeftBehindLampState_Fr	Ok.	04	00 00 00 00	[0F]

Fig. 98.9 Left-front Headlamp Open Load Message

Fig.98.10 LIN Network Message



98.4 Conclusion

LIN bus is a great option in the automotive light control system with low cost but high reliability by using the UART interface. Kinetis KEA has many I/O and communication sources and inherits excellent characteristics of ARM architecture, such as low-power consumption and high performance. This chapter chooses the LIN

bus as the communication network and KEA as the core microprocessor to develop an automotive light control system with the implementation of the light control and diagnosis functions. Further studies on new light control can take this system as a basis so that there is no need to build the LIN network and low-level firmware of Kinetis KEA64, which can greatly save developing time.

References

1. Giradkar M, Khanapurkar M. Design and implementation of adaptive front light system of vehicle using FPGA based LIN controller [C]. 4th international conference on emerging trends in engineering and technology (ICETET). IEEE; 2011. pp. 258–61.
2. LIN Consortium, 2010. LIN Specification Package Revision 2.2A[S]. December 31, 2010. www.lin-subbus.org.
3. Hocenski Z, Keser T, Nenadic K. Distributed intelligent control of car lighting system with fault detection [C]. Industrial electronics society (IECON). 33rd annual conference of the IEE. IEEE; 2007. pp. 498–503.
4. Xu YY, Wang JC, Chen WX, et al. Application of LIN bus in vehicle network [C]. IEEE international conference on vehicular electronics and safety (ICVES). IEEE; 2006. pp. 119–23.
5. KEA64 sub-family reference manual, Rev 1, 01[M]. Freescale Semiconductor, Inc. 2014.
6. Kum D, Park GM, Lee S, Jung W. AUTOSAR migration from existing automotive software [C]. International conference on control, automation and systems (ICCAS). IEEE; 2008. pp. 558–62.
7. Luo F, Lin YP. Design of the common framework for LIN low level driver [C]. 5th International conference on digital image processing (ICDIP). SPIE; 2013. pp. 1–10.
8. Lawrenz W. Communication protocol conformance testing—example LIN—[C]. Vehicular electronics and safety, ICVES 2006. IEEE; 2006. pp. 155–62.

Chapter 99
A Modified Design of an Active Magnetic Bearing Controller

Long He, Xiaoqiao Shi and Xiaolong Chen

Abstract This chapter mainly discusses electromagnetic bearing control theories and control algorithms. A modified design of a controller is presented to establish the closed-loop control system. By adopting the integrating element in a proportional integral derivative (PID) controller, the effects of variation of load on the position of the rotor can be compensated. By selecting the voltage control method instead of current control, the model of the system has better robustness. Finally, the simulation results are given to study the system responses of appropriate parameters to check the feasibility of the algorithm.

Keywords Electromagnetic bearing · Controller · System parameters

99.1 Introduction

First applications of the electromagnetic suspension principle have been in experimental physics, and suggestions to use this principle for suspending transportation vehicles for high-speed trains go back to 1937. There are various ways of designing magnetic suspensions for a contact free support—the active magnetic bearing (AMB) is just one of them [1]. AMB is a collection of electromagnets used to suspend an object via feedback control. Commercial applications include pumps, compressors, flywheels, milling and grinding spindles, turbine engines, and centrifuges [2].

L. He (✉)
The Aeronautic Maintenance Engineering Department, Chengdu Aeronautic Vocational and Technical College, 610100 Chengdu, China
e-mail: he.long@msn.com

X. Shi · X. Chen
The School of Electro-Mechanical Engineering, Xidian University, 710071 Xi'an, China
e-mail: shixqfcgz_90@126.com

X. Chen
e-mail: xlchen@mail.xidian.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_99

Magnetic suspension offers advantages over conventional bearings such as lower rotating losses and higher speeds. Bleuler developed a method for designing decentralized proportional derivative (PD) controllers for rigid rotors in 1984, and Salm introduced a control procedure for flexible rotors with collocation and gave stability guarantees for the case of continuous time control in 1988. H_∞/μ-synthesis control was first applied to AMB systems around 1990 [3]. A set of decoupled PD controllers was designed and the stability robustness was evaluated via μ-analysis in 2003 [4]. Fault detection and diagnosis were integrated into the active magnetic bearing system to provide reliable information about the system state in 2009 [5].

In the above AMB systems, the controller is one of the key techniques. The performance directly affects the work of magnetic bearing. So this chapter mainly describes the control theories and improves the algorithm.

This chapter shows the difference between the PD controller and proportional integral derivative (PID) controller. By adopting the integrating element in the PID controller, the effects of variation of load on the position of the rotor can be compensated. By selecting the voltage control method instead of current control, the model of the system has better robustness. At the end of the chapter, simulations of different control parameters are given to check the feasibility of the algorithm.

Figure 99.1 presents the main components of a control system of AMB (a controller, a power amplifier, and sensors). A sensor detects the displacement of the rotor from its reference position, and then the controller derives a control signal from the measurements. A power amplifier transforms this control signal into control current, and the control current generates a magnetic field in the actuating magnets, resulting in magnetic force. Then the rotor remains in its hovering position [6–8].

This chapter is devoted to the research and design of the PID controller of AMB. There are three parts: first, the design and selection of the control methods; second, the analysis of the structure and transfer function of the closed-loop system; and last, simulation and parameter selection.

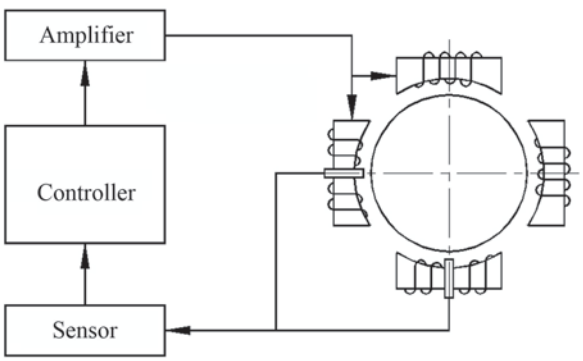


Fig. 99.1 Basic composition of active magnetic bearing

99.2 The Principle and Design of the PID Controller

As mentioned before, we discuss the relationships between force and displacement in the electromagnetic field of AMB. When the rotor deviates from the working position, the magnetizing current will change to alter the rotor force, making the rotor return to the operating point and then achieve controlling targets. In order to analyze the system, one direction of AMB controller is studied.

The transient force, a function of displacement and current, is a linear equation:

$$f(x,i) = k_x x + k_i i \tag{99.1}$$

The control of the rotor is to provide a restoring force, e.g., similar to the mechanical spring. Besides, a damping component should be provided to reduce oscillations. Then the support force f can be calculated as follows [1]:

$$f = -kx - d\dot{x} \tag{99.2}$$

where k is stiffness, d is damping, $\dot{x} = dx/dt$ is the rotor speed in the direction of this degree of freedom, force f and displacement x are both deviations from the equilibrium point.

According to Eqs. (99.1) and (99.2), the control current $i(x)$ is as follows:

$$i(x) = -\frac{(k + k_x)x + d\dot{x}}{k_i} \tag{99.3}$$

Here, let

$$k_p = -(k + k_x)/k_i \tag{99.4}$$

$$k_d = -d/k_i \tag{99.5}$$

and then

$$i(x) = k_p x + k_d \dot{x} \tag{99.6}$$

Such a control law is named PD control. For an active magnetic bearing with PD control, an external static load will always result in a change in the steady position which is undesired in a technical application.

Thus, an integral part is needed to be introduced to compensate the effect. The inductance L varies in different rotor positions. The rotor motion also generates a coil voltage. This induced voltage is [3]

$$u = Ri + \frac{Ldi}{dt} + k_u \dot{x} \tag{99.7}$$

where R is the coil resistance and k_u is the voltage—the speed factor.

Equation (99.7) is the voltage control; this method maintains the stability of a rotor through a change in voltage, resulting from the signal form sensors.

Figure 99.2 shows the control system diagram of the voltage control scheme.

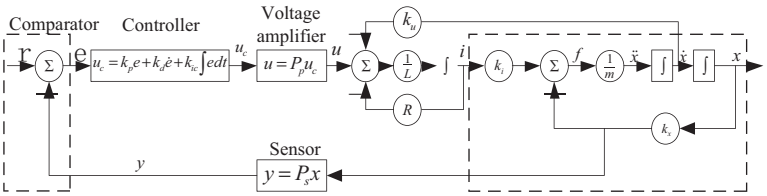


Fig. 99.2 System diagram of the voltage control scheme

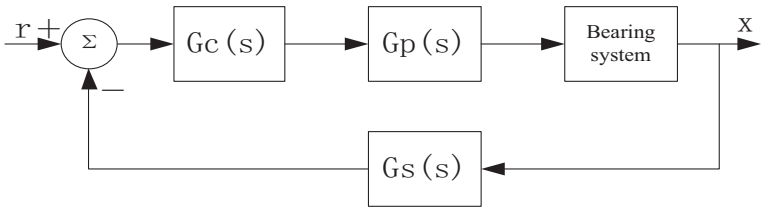


Fig. 99.3 System block diagram of the closed-loop control system

The system block diagram is shown in Fig. 99.3. $G_s(s)$ is the transfer function of the position sensor, $G_c(s)$ is the transfer function of the PID controller, and $G_p(s)$ is the transfer function of the amplifier.

$$G_c(s) = k_p + \frac{k_{ic}}{s} + k_d s \tag{99.8}$$

where k_p, k_{ic} and k_d are the proportional gain coefficients, the integral gain coefficients, and the differential gain coefficients, respectively.

$$G_p(s) = \frac{\lambda}{1 + sT_p} \tag{99.9}$$

$$G_s(s) = \frac{A_s}{1 + sT_s} \tag{99.10}$$

where λ is the gain factor of the amplifier circuit, T_p is the decay time constant of the amplifier circuit, A_s is the gain factor of the displacement sensor, and T_s is the decay time constant of the displacement sensor.

As mentioned above, PID control can compensate the effects of load changes on the steady-state position of the rotor, so this chapter selects PID control.

The motion equation of controlled objects is

$$m\ddot{x} = k_x x + k_i i \tag{99.11}$$

Combining the formulas above, we get the equation of the system

$$ms^3 + Kk_d s^2 + (Kk_p - k_x)s + Kk_{ic} = 0 \tag{99.12}$$

From the above analysis, PID control introduces an integral part, which can compensate for the effects of variation of load on the steady-state position of the rotor. So, PID control is better than PD control at this extent. The voltage control method considers the prevention of magnetic bearing coil inductance on the current changes compared to the current control. It makes the mathematical model of the system more accurate and has better robustness. Besides, the open-loop instability of the system improves. The downsides of the structure are that it is relatively complex, and the controller requires at least two stages. So, for a small-scale system, the current control is available.

99.3 The Simulation Results

The parameters k_p,k_{ic} and k_d are determined by stiffness k and damping d in the closed-loop system. In this chapter, MATLAB is used to simulate conditions of a PID controller. We can study the system response of different gain coefficients.

The control function of the PID controller is

$$u(t) = k_p e + k_{ic} \int_0^t e dt + k_d \frac{de}{dt} \tag{99.13}$$

$$e = r - u_1(t) \tag{99.14}$$

where e is the input of the controller, r is a reference value or control objective, and $u_1(t)$ is the output of the sensor.

The transfer function of the PID controller is

$$G_c(s) = k_p + \frac{k_{ic}}{s} + k_d s \tag{99.15}$$

The transfer function of the closed-loop system simulated is

$$H(s) = \frac{G_c(s)G_p(s)G(s)}{1 + G_c(s)G_p(s)G(s)G_s(s)} \tag{99.16}$$

Put Eqs. (99.9) and (99.10) in Eq. (99.16) and we get:

$$H(s) = \frac{k_i \lambda \left(k_d s^2 + k_p s + k_{ic}\right)(T_s s + 1)}{s\left(ms^2 - k_x\right)\left(T_p s + 1\right)\left(T_s s + 1\right) + k_i \lambda A_s \left(k_d s^2 + k_p s + k_{ic}\right)} \tag{99.17}$$

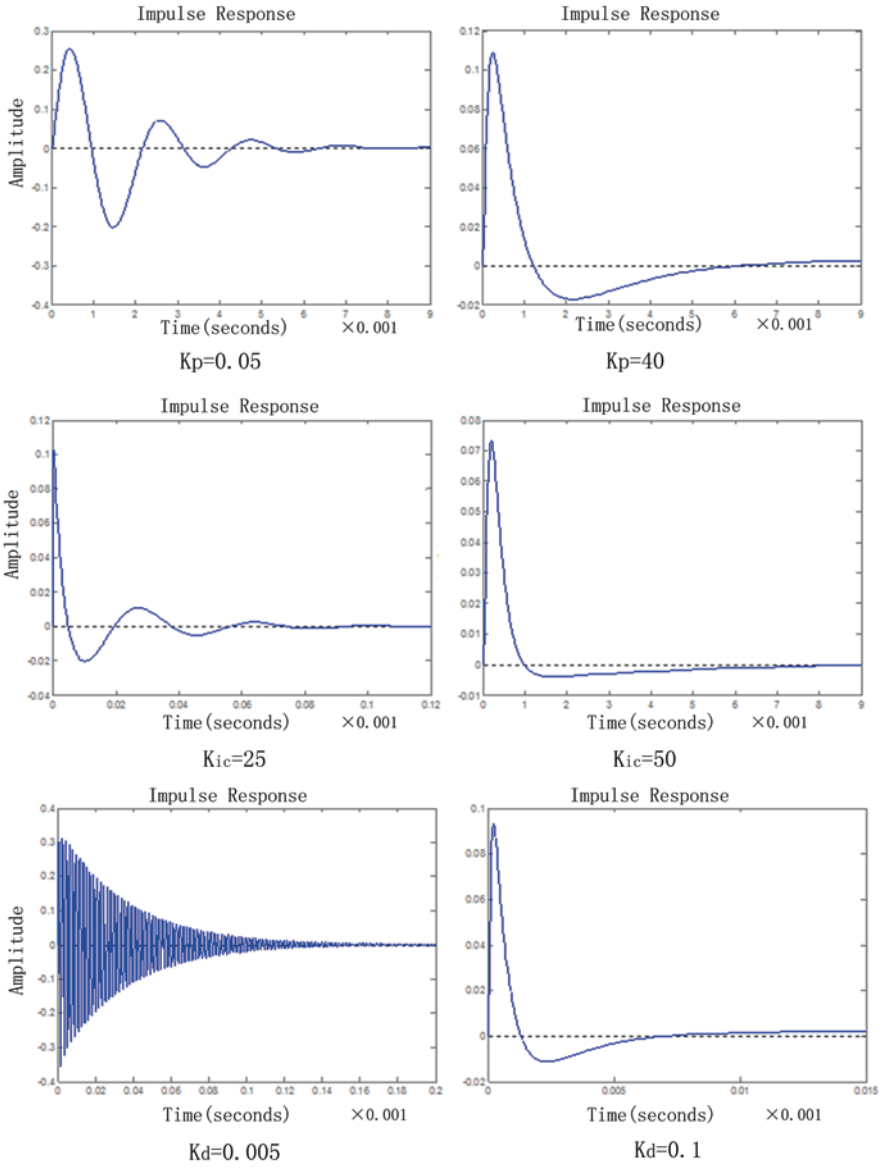


Fig. 99.4 Closed-loop response graphs of different PID parameters of the system

The relevant parameters used in the simulation are: the quality of the rotor $m = 4.5kg$, the bias current of the coil $I_0 = 0.4A$, the displacement stiffness coefficient $k_x = 4.211\times10^5\, N/m$, the current stiffness coefficient $k_i = 157.9N/A$, the gain coefficient of the sensor $A_s = 10^4$, the time constant $T_s = 7.958\times10^{-5}\, s$, the gain factor of the power amplifier $\lambda = 1$, and time constant $T_p = 3.183\times10^{-5}\, s$.

Figure 99.4 shows the simulation results using the parameters above, with different parameters, k_p,k_{ic} and k_d.

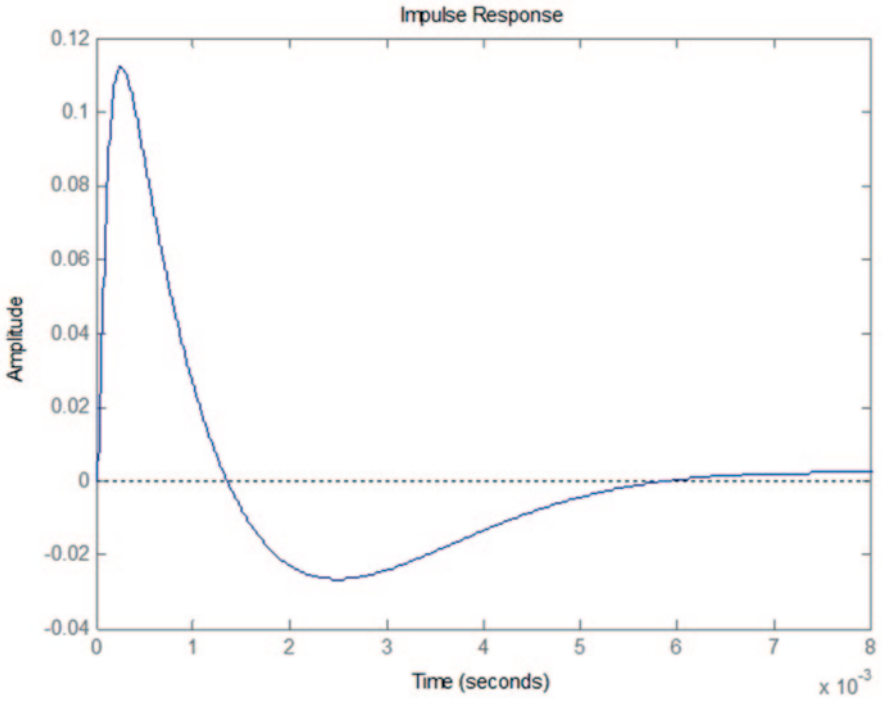


Fig. 99.5 Closed-loop response of the system

The simulation results show that k_p has the greatest impact on the dynamic performance. When k_p is small, the system will oscillate at the beginning. With the increase of k_p, the time of convergence increases. So the best range of k_p is 1–15. k_{ic} has no significant effect at the start, but the time of convergence is significantly reduced when it increases. So, the optimum range is 20–50. k_d has little effect on initial displacement. As k_d increases, the time of convergence increases. Obviously, when k_d is less than 0.001, the system undergoes an oscillation phenomenon. So, the best range of k_d is 0.01–0.1.

Through the analysis above, we carry out multigroup experiments and choose $k_p = 5, k_{ic} = 40$ and $k_d = 0.05$. Fig. 99.5 shows the closed-loop impulse response.

99.4 Conclusion

A modified AMB control algorithm is designed in this chapter. An integral part is introduced to compensate for the effect of variation of load on the position of the rotor. The voltage control method is adopted to improve the performance of transients of load. Through simulations, the effect of control parameters on the closed-loop response of the system indicates the feasibility of this algorithm.

References

1. Schwertzer G, Maslen EH. Magnetic bearings theory, design, and application to rotating machinery. Berlin: Springer; 2009. pp. 5–55.
2. Lin CT, Jo CP. GA-based fuzzy reinforcement learning for control of a magnetic bearing system. IEEE T Syst Man Cyb. 2000;30(2):276–89.
3. Chen J, Liu K, Xiao K. H∞ control of active magnetic bearings using closed loop identification model. IEEE international conference on mechatronics and automation. Los Alamitos: IEEE Computer Society. 2011. pp. 349–53.
4. Stephens LS, Chin H-M. Robust stability of the Lorentz-type self bearing servomotor. Proceedings of the eighth international symposium on magnetic bearings. JSME Int J Ser C. 2003;46(2):355–62.
5. Beckerle P, Butzek N, et al. Application of a balancing filter for model-based fault diagnosis on a centrifugal pump in active magnetic bearings. ASME international design engineering technical conferences and computers and information in engineering conferences. American society of mechanical engineers. vol. 3. New York; 2009. pp. 215–222.
6. Hong S-K, Langari R. Robust fuzzy control of a magnetic bearing system subject to harmonic disturbances. IEEE T Contr Syst T. 2000;8(2):366–71.
7. Zhang W, Yefa H. A prototype of flywheel energy storage system suspended by active magnetic bearings with PID controller. APPEEC–proceedings. Institute of electrical and electronics engineers computer society. Piscataway; 2009. pp. 1–4.
8. Wang X, Jiang K. Study on the centripetal effect of the magnetic bearing. Electrical and control engineering (ICECE). Los Alamitos: IEEE Computer Society; 2010. pp. 2135–8.

Chapter 100
Construction of a One-Coincidence Frequency-Hopping Sequence Set with Optimal Performance

Hong Wang and Ping Huang

Abstract In multiple-access spread-spectrum communication systems and multiuser radar systems, it is important to construct an optimal frequency-hopping sequence set to improve the efficiency of communication. By the Cartesian product, one class of one-coincidence frequency-hopping sequence set is constructed for the first time. It shows that this set has optimal balance properties and optimal average periodic Hamming correlation according to the respective bounds in theory and experiment. This method is very simple, feasible, and has better applied foreground.

Keywords Cartesian product · Frequency-hopping sequence · One-coincidence frequency-hopping sequence

100.1 Introduction

Due to the antijamming, secure and multiple access properties, frequency-hopping multiple accesses (FHMA) spread spectrum systems have been applied to Bluetooth and sonar echolocation systems where each user is denoted by a frequency-hopping sequence (FHS) [1–3]. An important demand is to keep the mutual interference between transmitters as low as possible. It is obvious that the level of the mutual interference mainly depends on the Hamming correlation properties of the FHSs [1–4]; therefore, it is important to construct an optimal FHS or FHS set to improve the efficiency of communication.

The concept of the average periodic Hamming correlation (APHC) was presented by Peng [5, 6]. APHC shows the global average error rate of the frequency-hopping communication system, which is another measure used to estimate the optimal

H. Wang (✉) · P. Huang
Chengdu Neusoft university, 618201 Chengdu, China
e-mail: cdyjq14@163.com

P. Huang
e-mail: 153522214@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_100

properties of a FHS set. According to the Peng-Fan bound [7], a FHS set is optimal; but it may not be optimal according to the APHC bound [5]. To improve the efficiency of communication, it is important to construct an optimal FHS set according to the Peng-Fan bound [7] and the APHC bound [5].

In this chapter, based on the Cartesian product, we present a new class of one-coincidence FHS sets with optimal performance. For the remaining content of this chapter, we have made the following arrangement: In Sect. 100.2, we introduce notations and definitions; then, we present a new generic construction for a one-coincidence FHS set with optimal performance in Sect. 100.3; finally, we conclude the chapter in Sect. 100.4.

100.2 Preliminaries

Let p frequency slots form a frequency slot set $F=\{f_1,f_2,\ldots,f_p\}$. Assume S is a set length N which is defined over F. For any two FHSs $y=(y_0,y_1,\ldots,y_N-1), z=(z_0,z_1,\ldots,z_N-1)\in S$, the periodic Hamming correlation function $H_{y,z}(\tau)$ of x and z at time delay τ is defined as below:

$$H_{y,z}(\tau)=\sum_{k=0}^{N-1} h(y_k,z_{k+\tau}) \quad (0\le\tau<N). \tag{100.1}$$

$h(y_k,z_{k+\tau})=1$ if $y_k=z_{k+\tau}$, $h(y_k,z_{k+\tau})=0$ otherwise, where all operations are performed modulo N among the position indices.

The maximum periodic Hamming autocorrelation $H_a(S)$ is defined as follows:

$$H_a(S)=\max_{1\le\tau<N}\left\{H_{y,y}(\tau)\,|\,y\in S\right\}. \tag{100.2}$$

The maximum periodic Hamming cross-correlation $H_c(S)$ is defined as below:

$$H_c(S)=\max_{0\le\tau<N}\left\{H_{y,z}(\tau)\,|\,y,z\in S, y\ne z\right\}. \tag{100.3}$$

The maximum periodic Hamming correlation $H_m(S)$ is defined as below:

$$H_m(S)=\max\left\{H_a(S),H_c(S)\right\}. \tag{100.4}$$

Peng and Fan [7] developed the lower bound on H_m.

Lemma 1 Assume S is a FHS set with the length N of size M and over a frequency slot set F with size p, then

$$H_m\ge\left\lceil\frac{(NM-q)N}{(NM-1)q}\right\rceil. \tag{100.5}$$

This bound is called Peng-Fan bound, S is said to be optimal if the equality in (100.5) is achieved.

There are two integers l_a and l_c with $l_a>0$, $l_c>0$. Then, the low-hit zone L_H of S is defined as below:

$$L_H = \min\left\{ \max_{\substack{0<\tau\le T \\ \forall y\in S}} \left\{T \,|\, H_{y,y}(\tau) \le l_a\right\}, \max_{\substack{0\le\tau\le T \\ \forall y,z\in S, y\ne z}} \left\{T \,|\, H_{y,z}(\tau) \le l_c\right\} \right\}. \quad (100.6)$$

If $L_H>0$ holds, S is called the LHZ FHS set. For the LHZ FHS set, in 2006, Peng, Fan and Lee [8] obtained the bound called Peng-Fan-Lee bound.

Lemma 2 Assume S is an FHS set with length N of size M over a frequency slot set F with size p, and L the LHZ of S according to constant H_m. Then, for the integer Z, $1 \le Z \le L$, we obtain

$$H_m \ge \left\lceil \frac{(NZ+M-q)N}{(MZ+M-1)q} \right\rceil. \quad (100.7)$$

By the Peng-Fan-Lee bound (100.7), S is an optimal set if the inequality (100.7) holds with equality.

What is more, the APHC is another important performance indicator of the FHS set.

Definition 1 ([5]) Let S be an FHS set with length N of size M over a frequency slot set F with size v. Then, the overall number of the periodic Hamming autocorrelation is defined as follows:

$$S_a(s) = \sum_{\forall y\in S, 1\le\tau\le N-1} H_y(\tau). \quad (100.8)$$

And the periodic Hamming cross-correlation is defined as below:

$$S_c(s) = \frac{1}{2} \sum_{\forall y,z\in S, y\ne z, 0\le\tau\le N-1} H_{y,z}(\tau). \quad (100.9)$$

The average periodic Hamming autocorrelation is defined as below:

$$A_a(s) = \frac{S_a(s)}{M(N-1)}. \quad (100.10)$$

The average periodic Hamming cross-correlation is given as below:

$$A_c(s) = \frac{2\,S_c(s)}{MN(M-1)}. \quad (100.11)$$

For the parameters v, N, M, $A_a(s)$ and $A_c(s)$ of S, Peng et al. [6] presented the following bound.

Lemma 3 (APHC bound [6]) Provided that S is a FHS set of size M with length N over the frequency slot set F with size v, $A_a(s)$ and $A_c(s)$ are the average periodic Hamming autocorrelation and the average periodic hamming cross-correlation of S, then

$$\frac{A_a(s)}{N(M-1)}+\frac{A_c(s)}{N-1}\geq\frac{MN-v}{v(N-1)(M-1)}. \tag{100.12}$$

Hereafter, we use the following definition:

A FHS set S is an optimal FHS set according to the APHC bound if $A_a(s)$ and $A_c(s)$ of S satisfy inequality (100.12) with equality.

In 2010, Peng [9] constructed several FHS sets with optimal APHC which are based on cubic polynomials.

100.3 Construction of an Optimal FHS Set

The Cartesian product of k nonempty sets $A_0, A_1,\cdots, A_{k-1}$ is the gather of all ordered pairs of the elements

$$A_0\times A_1\times\cdots\times A_{k-1}=\left\{\left(x_0,x_1,\cdots,x_{k-1}\right): x_0\in A_0, x_1\in A_1,\cdots,x_{k-1}\in A_{k-1}\right\}. \tag{100.13}$$

Like the set's operations, the Cartesian product operation is used in a series of sequences. If $\left\{s_0(t)\right\}_{t=0}^{N_0-1}$, $\left\{s_1(t)\right\}_{t=0}^{N_1-1},\cdots,\left\{s_{k-1}(t)\right\}_{t=0}^{N_{k-1}-1}$ is k sequences of periods $N_0, N_1,\cdots,$and N_{k-1} over $F_0, F_1,\cdots,$and F_{k-1}, then their Cartesian product results in a set of group of ordered elements, each group of ordered elements comprises of "k" elements, let $N=\text{lcm}(N_0, N_1,\cdots, N_{k-1})$, it can be stated as

$$\left\{(s_0(0), s_1(0),\cdots, s_{k-1}(0)),\cdots,(s_0(t), s_1(t),\cdots, s_{k-1}(t))\right\}. \tag{100.14}$$

It is conspicuous that the Cartesian produces the above results in a k-dimensional vector sequence. J. H. Chung [10] gave an universal method to construct the LHZ FHS set from known FHS sets by the Cartesian product for the first time. In this section, we construct a new class of one-coincidence FHS set with the optimal Hamming correlation. There are some comparisons between the existing FHS sets and the new one is listed in Table 100.1.

Table 100.1 Known and new FHS set

Method	Constructions	Peng-Fan-Lee bound	Peng-Fan bound	APHC bound	Balance property
Cartesian product	Theorem 1 in [10]	Optimal	–	Optimal	It is difficult to determine
Cartesian product	Theorem 2 in [10]	Optimal	–	Not optimal	
Cartesian product	Theorem 3 in [10]	Optimal	–	Not optimal	
Cartesian product	Theorem 4 in [10]	Optimal	–	Not optimal	
Cyclostomes theory	Theorem 2 in [4]	–	Optimal	Not optimal	
Cartesian product	Theorem 1 here	–	Optimal	Optimal	Good

100.3.1 Construction A

Step1 Select a FHS set $X=\{x_i(t)\}_{t=0}^{p-2}$ with parameters $(p-1,p,p,1)$as the base FHS set which defines over the alphabet set F_0 with a size of p. For any two FHSs x_{l_0}, x_{l_1}in X, the periodic Hamming correlation is given by

$$H_{x_{l_0},x_{l_1}}(\tau)=\begin{cases}p-1 & \text{if } \tau \bmod (p-1)\equiv 0 \text{ and } l_0=l_1\\ 0 & \text{if } \tau \bmod (p-1)\neq 0 \text{ and } l_0=l_1\\ 0 & \text{if } \tau \bmod (p-1)\equiv 0 \text{ and } l_0\neq l_1\\ 1 & \text{otherwise}\end{cases} \tag{100.15}$$

select another FHS set $\{y_k(t)\}_{t=0}^{q-1}$ as the base FHS set Y with parameters $(q,q-1,q,1)$; as to any two FHSs y_i, $y_j$in Y, such that

$$H_{y_i,y_j}(\tau)=\begin{cases}q & \text{if } \tau\equiv 0 \bmod q \text{ and } i=j\\ 0 & \text{if } \tau\neq 0 \bmod q \text{ and } i=j\\ 1 & \text{otherwise}\end{cases} \tag{100.16}$$

$\gcd(p-1,q)=1$, let $\min(p,q-1)=p$. Select any p different sequences from the base sequence set Y to form the FHS set Z.

Step2 Define a new FHS set S over $F_0\times F_1$ from Z and X as below:

$$\begin{cases}S=\{S_{i_0}(t)\mid 0\le t<(p-1)q\}\\ S_{i_0}(t)=(X_{i_0}(t \bmod (p-1)),\ Z_{i_0}(t \bmod q))\ \ 0\le i_0<p\end{cases} \tag{100.17}$$

Theorem 1 S is a one-coincidence FHS set with parameterse $((p-1)q, p, qp, 1)$.

Proof: as to any sequences S_{i_0} and $S_{i_0'}$ in the FHS set S, by the definition of periodic Hamming correlation, we obtain

$$\begin{aligned} H_{S_{i_0}, S_{i_0'}}(\tau) &= \sum_{t=0}^{(p-1)q-1} h\left(S_{i_0}(t), S_{i_0'}(t+\tau)\right) = \sum_{t=0}^{(p-1)q-1} h\left(S_{i_0}(t), S_{i_0'}(t+\tau)\right) \\ &= \sum_{t=0}^{(p-1)q-1} h\Big(\left(x_{i_0}(t), z_{i_0}(t)\right), \left(x_{i_0'}(t+\tau), z_{i_0'}(t+\tau)\right)\Big) \\ &= \sum_{t=0}^{(p-1)q-1} h\Big(\left(x_{i_0}(t), x_{i_0'}(t+\tau)\right), \left(z_{i_0}(t), z_{i_0'}(t+\tau)\right)\Big) \\ &= \sum_{t_0=0}^{p-2} h\left(x_{i_0}(t_0), x_{i_0'}(t_0+\tau)\right) \cdot \sum_{t_1=0}^{q-1} h\left(z_{i_0}(t_0 + t_1(p-1)), z_{i_0'}(t_0 + t_1(p-1)+\tau)\right). \end{aligned} \tag{100.18}$$

As to any t_1, $0 \le t_1 < q$, t_0, $t_0 + p - 1$, $t_0 + 2(p-1)$, … , $t_0 + (q-1)(p-1)$ is q different positive integers, so

$$H_{S_{i_0}, S_{i_0'}}(\tau) = \sum_{t_0=0}^{p-2} h(x_{i_0}(t_0), x_{i_0'}(t_0+\tau)) \cdot \sum_{t_1=0}^{q-1} h(z_{i_0}(t_1), z_{i_0'}(t_1+\tau)). \tag{100.19}$$

Therefore, the maximum periodic Hamming autocorrelation of S be given by when $i_0 = i_0'$

$$H_a(S) = 0. \tag{100.20}$$

For any i_0 and i_0', $0 \le i_0 \ne i_0' < p$, the following equation may not hold:

$$i_0 \equiv i_0' \mod p. \tag{100.21}$$

Thus, the maximum periodic Hamming cross-correlation of S can be given by

$$H_c(S) = \begin{cases} 0 & \text{if } \tau \equiv 0 \mod (p-1) \\ 1 & \text{otherwise} \end{cases} \tag{100.22}$$

The maximum periodic Hamming correlation of S is expressed by

$$H_m(S) = 1. \tag{100.23}$$

The statement holds and the proof is completed.

In addition, it is easy to verify S is optimal according to the Peng-Fan bound (100.5).

100.4 Performance Analysis

(1) Balance property The balance property of the FHS set is measured according to the parameter σ:

$$\sigma = \frac{q}{N}\sqrt{\frac{1}{q}\sum_{i=0}^{q-1}\left(N_{f_i} - \frac{N}{q}\right)^2} \tag{100.24}$$

where N_{f_i} denotes the number of the frequency slot f_i which appears in a sequence; the parameter σ shows the uniformity of each frequency slot in sequence; in the ideal case, $N_{f_i} = \frac{N}{q}$, that is $\sigma = 0$. The smaller the value of σ is, the better the balance properties of S will be. The value of σ of S can be given by

$$\sigma = \frac{qp}{(p-1)q}\sqrt{\frac{1}{qp}\sum_{i=0}^{qp-1}\left(1 - \frac{(p-1)q}{qp}\right)^2} = \frac{1}{p-1}. \tag{100.25}$$

The above discussions have shown that S has good balance properties.

(2) APHC property As proved by theorem 1, the distribution of the periodic Hamming correlation of any sequences S_{i_0} and $S_{i_0'}$ in the FHS set S is given by

$$H_{S_{i_0}, S_{i_0'}}(\tau) = \begin{cases} 0 & \text{if } 0 < \tau \le (p-1)q-1, \quad i_0 = i_0' \\ 0 & \text{if } \tau \equiv 0 \bmod (p-1), \quad i_0 \ne i_0' \\ 1 & \text{if } 0 \le \tau \le (p-1)q-1, \quad i_0 \ne i_0' \end{cases} \tag{100.26}$$

Now, we check the optimal properties of S.

$$\begin{aligned} &S_a(s) = 0,\ A_a(s) = 0 \\ &2\,S_c(s) = (p-1)qp(p-1) - qp(p-1) = (p-2)qp(p-1) \\ &\frac{A_a(s)}{N(M-1)} + \frac{A_c(s)}{N-1} = \frac{(p-2)qp(p-1)}{p(p-1)q(p-1)((p-1)q-1)} = \frac{p-2}{(p-1)((p-1)q-1)} \\ &\frac{MN-v}{v(N-1)(M-1)} = \frac{p(p-1)q - pq}{pq((p-1)q-1)(p-1)} = \frac{p-2}{(p-1)((p-1)q-1)} \end{aligned} \tag{100.27}$$

It is obvious that the equality holds in the inequality (100.12), therefore, the FHS set S is optimal according to the APHC bound (100.12).

The following example can be used to illustrate the construction.
Let $p=31$, we construct the one-coincidence FHS set X of length 31 as follows:
$X_0=\{0, 10, 20, 30, 9, 19, 29, 8, 18, 28, 7\ldots\}$; $X_1=\{0, 11, 22, 2, 13, 24, 4, 15, 26, 6, 17\ldots\}$;
Let $q=7$, we construct the one-coincidence FHS set Y of length 6 as follows:
$Y_0=\{1, 3, 2, 6, 4, 5\}$; $Y_1=\{2, 4, 3, 0, 5, 6\}$; $Y_2=\{3, 5, 4, 1, 6, 0\}$; $Y_3=\{4, 6, 5, 2, 0, 1\}$;
By applying construction A, we obtain an FHS set with length 186 as follows:
$S_0=\{(1, 3), (2, 3), (3, 6), (4, 4), (5, 5), (6, 1), (7, 3), (8, 2), (9, 6), (10, 4), (2, 2), \ldots\}$
$S_6=\{(7, 2), (14, 8), (21, 5), (28, 3), (4, 4), (11, 0), (18, 2), (25, 1), (1, 5), (15, 4) \ldots\}$.

The maximum periodic Hamming autocorrelation of S is 0, the maximum periodic Hamming cross-correlation of S is shown in Fig. 100.1.

The total number of the periodic Hamming autocorrelation and the periodic Hamming cross-correlation of S is $A_a=0$ and $2\,S_c(s)=6510$, respectively. The APHC of S can be given by:

$$A_c(s)=\frac{2\,S_c(s)}{MN(M-1)}=\frac{6510}{7\times31\times6\times6}=\frac{5}{6}$$

By putting $N=186$, $M=7$, $A_a(S)=0$, and $A_c(s)=\frac{5}{6}$ in (100.12), we can see that both sides of (100.12) become $\frac{1}{222}$. Thus S is optimal according to the APHC bound (100.12). In addition, it is easy to verify that S is optimal according to the Peng-Fan bound (100.5); furthermore, the parameter $\sigma=\frac{1}{6}$; thus S has good balance properties.

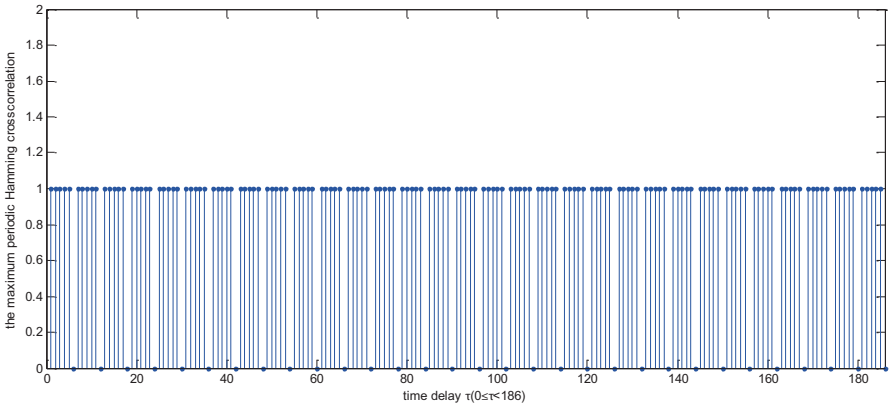


Fig. 100.1 Maximum periodic Hamming cross-correlation and correlation of S

100.5 Conclusion

To improve the efficiency of the frequency-hopping communication system, it is important to construct an optimal FHS or FHS set. On this basis, we construct one class of one-coincidence FHS set by the Cartesian product. It is a challenging work to construct other classes of FHS sets by the Cartesian product.

References

1. Fan PZ, Darnell M. Sequence design for communications applications [M]. London: Research Studies Press, Wiley; 1996. pp. 1–20.
2. Golomb SW, Gong G. Signal design for good correlation: for wireless communication, cryptography and radar [M]. Cambridge: Cambridge University Press; 2005. pp. 2–40.
3. Lempel A, Greenberger H. Families of sequences with optimal Hamming correlation properties [J]. IEEE Trans Inform Theory. 1974;20(1):90–4.
4. Ren WL, Fu FW, Zhou ZHCH. New sets of frequency-hopping sequences with optimal Hamming correlation [J]. Des Codes Cryptogr. 2014;72:423–34.
5. Peng DY, Peng T, Fan PZ. Generalized class of cubic frequency-hopping sequences with large family size [J]. IEE Proc Commun. 2005;15(2):897–902.
6. Peng DY, Peng T, Tang XH. A class of optimal frequency hopping sequences based upon the theory of power residues [C]. SETA 2008, Proceedings of the 5th international conference on Sequences and Their Applications [C]. 2008; Springer, Germany, 5203:188–196.
7. Peng DY, Fan PZ. Lower bounds on Hamming auto and cross correlations of frequency hopping sequences [J]. IEEE Trans Inform Theory. 2004;50(12):2149–4.
8. Peng DY, Fan PZ. Lower bounds on the Hamming auto and cross correlations of frequency-hopping sequences [J]. IEEE Trans Inform Theory. 2004;50(9):2149–4.
9. Peng DY, Niu X, Tang X. Average hamming correlation for the cubic polynomial hopping sequences [J]. IET Commun. 2010;4(15):1775–86.
10. Chung JH, Yang K. New classes of optimal low-hit-zone frequency-hopping sequence sets by Cartesian product [J]. IEEE Trans Inform Theory. 2013;59(1):726–32.

Chapter 101
Universal Control of Second-Order Nonholonomic Systems

Xu Shen, Jian Zhang and Mingjun Li

Abstract Three global universal controllers based on dynamic feedback techniques are proposed for a second-order nonholonomic system (NS). A Lyapunov function is defined for the closed-loop system to guarantee exponential convergence. If the initial state is limited, K-exponential stability is obtained; however, the error states and the control inputs will go to frequent oscillation with peaking phenomena. In order to avoid this problem, the control objective is loosened to practical stability. As the three global controllers need no persistent excitation condition (PEC), they can be simultaneously applied to the point stabilization and the trajectory tracking of NS. The simulation results show the effectiveness of the proposed controllers.

Keywords Nonholonomic systems · Global universal controller · Point stabilization · Trajectory tracking · K-exponential stability

101.1 Introduction

The nonholonomic system (NS) has been the focus in the control domain [1] because there is no continuous static feedback control law to stabilize it [2]. The NS has two tasks, i.e., the trajectory tracking and the point stabilization, which cannot imply each other. Trajectory tracking under persistent excitation conditions (PECs) is not limited [3] by the constraint [2], and classical methods are available, e.g., the approximate linearization along the target track [4], PID [5], etc.; thus, the PEC has been almost a standard assumption in the literature; however, the PEC excludes the point stabilization. Most driftless systems have no controllable linearized forms without the PEC [6]. As a result, most papers deal with the two aspects separately.

M. Li (✉) · J. Zhang
Department of Urban Rail Transportation, Beijing Vocational College of Transportation, 100096 Beijing, China
e-mail: lmjlmj8@163.com

X. Shen
Thermal Power Plant, Daqing Petro-chemical Branch Company, 163714 Daqing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_101

In the electrical and electronic engineering, many industrial robots (e.g., robots in the automobile assembly workshop) can be modeled in the NS. Their control usually involves two tasks; thus, it is necessary to study the universal control applicable to them [7]. The universal controller can be classified as below: (1) The one based on switching the two tasks. In the case of switching, the controllers need to reset parameters and restart [8–11], which thus causes oscillation; (2) the continuous time-varying universal controller. Controllers are not exponential convergent [7, 12]. Controllers limit the reference input to be a certain exponential attenuation signal [13, 14]; (3) the dynamic feedback universal controller. The one still needs switch when the tasks change [15]. The local asymptotic tracking controller [16] requires the target point to be in a reference trajectory with the PEC when used for point stabilization; (4) the practical stable universal controller. The typical design is the transversal function method [6, 17] based on Lie groups; however, it is easy to cause saturation and oscillation.

This chapter designs three global exponential universal controllers for a second-order NS. Compared with the above controllers, the controllers here feature: (1) No resetting parameters, (2) no PEC, (3) no saturation and oscillation, and (4) global K-exponential stability.

To simplify the writing, $\kappa(\cdot) \in FK(\chi)$ denotes that $\kappa(\cdot)$ is a class-K function of $\chi, \gamma(t) \in EXP(k)$ denotes that $\gamma(t)$ tends to be 0 exponentially with rate k.

101.2 Problem Formulation

Lemma 1 $\forall a_2 > a_1 > 0, m \geq 0, \exists c > 0, b \in [a_1, a_2) \Rightarrow t^m \, \mathrm{e}^{-a_2 t} \leq c \, \mathrm{e}^{-bt} (\forall t \geq 0)$.

Proof $\forall b \in [a_1, a_2)$, define $f(t) = t^m \mathrm{e}^{-(a_2 - b)t}. \frac{\mathrm{d}f(t)}{\mathrm{d}t} = 0 \Rightarrow t^* = \frac{m}{a_2 - b} \Rightarrow f(t) \leq f(t^*) \triangleq c \Rightarrow t^m \, \mathrm{e}^{-a_2 t} \leq c \, \mathrm{e}^{-bt} (\forall t \geq 0)$.

Lemma 2 [18] For system (1), $x \in \mathbb{R}^n, u \in \mathbb{R}^m, (\boldsymbol{A}, \boldsymbol{B})$ are controllable,

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \tilde{\boldsymbol{A}}(t)\boldsymbol{x} + \boldsymbol{f}(t) + \boldsymbol{B}\boldsymbol{u}, \tag{101.1}$$

where $\lim_{t \to \infty} \tilde{\boldsymbol{A}}(t) = 0, \int_0^\infty ||\tilde{\boldsymbol{A}}(t)|| \mathrm{d}t \leq c_1, ||\boldsymbol{f}(t)|| \leq c_2 \, \mathrm{e}^{-\lambda_0 t}, \lambda_0, c_2 \in (0, \infty), c_1 \in (0, \infty)$ or c_1 is a nonnegative nondecreasing function of $||\mathbf{x}(0)||$, there is:

1. $\boldsymbol{u} = -\mathbf{K}\mathbf{x}$ can be found so that (101.1) is an exponential convergent.
2. If $c_2 \in FK(\| \mathbf{x}(0) \|)$, then (101.1) is K-exponentially stable.

Consider the following second-order NS as its reference model:

$$\begin{cases} \ddot{x}_1 = u_1, \ddot{x}_2 = u_2, \ddot{x}_3 = x_2 u_1 \\ \ddot{x}_{1r} = u_{1r}, \dot{x}_{2r} = \overline{u}_{2r}, \ddot{x}_{3r} = x_{2r} u_{1r}, \end{cases} \tag{101.2}$$

where (u_1, u_2) are the inputs, $\mathbf{X} = [x_1, \dot{x}_1, x_2, \dot{x}_2, x_3, \dot{x}_3]^{\mathrm{T}}$ are the states, $(u_{1r}, \overline{u}_{2r})$ are the reference inputs, and $\mathbf{X}_r = [x_{1r}, \dot{x}_{1r}, x_{2r}, x_{3r}, \dot{x}_{3r}]^{\mathrm{T}}$ are the reference states.

Hypothesis 1 As to (101.2), a positive constant M exists so that
$\max(|u_{1r}(t)|, |u_{2r}(t)|, |x_{2r}(t)|, |\overline{u}_{2r}(t)|, |\dot{\overline{u}}_{2r}(t)|) \le M \cdot$
This hypothesis is not strict. Many real systems satisfy it. Define the error vector $\mathbf{X}_e = [x_{11}, x_{12}, x_{21}, x_{31}, x_{32}]^{\mathrm{T}}$, where

$$x_{11} = x_1 - x_{1r}, x_{12} = \dot{x}_1 - \dot{x}_{1r}, x_{21} = x_2 - x_{2r}, x_{31} = x_3 - x_{3r}, x_{32} = \dot{x}_3 - \dot{x}_{3r} \tag{101.3}$$

By (101.2), the error system can be written as

$$\begin{cases} \dot{x}_{11} = x_{12}, & \dot{x}_{21} = \overline{u}_2 - \overline{u}_{2r}, & \dot{x}_{31} = x_{32}, \\ \dot{x}_{12} = u_1 - u_{1r} & \dot{\overline{u}}_2 = u_2 \, \dot{x}_{32} = x_2 & u_1 - x_{2r} u_{1r} \end{cases}, \tag{101.4}$$

where $\overline{u}_2 = \dot{x}_2$. The tracking problem of (101.2) is: given Hypothesis 1 and bounded initial states, design control (u_1, u_2) so that $\dot{x}_2$ is bounded and $\mathbf{X}_{\mathrm{e}}$ tends to be 0 exponentially. If $x_{1r}(0), x_{2r}(0), x_{3r}(0)$ are the constants and $(u_{1r}(t), \overline{u}_{2r}(t)) \equiv 0 (\forall t \ge 0)$, then trajectory tracking degrades into point stabilization.

Introduce the variables as

$$y_1 = x_{12} + k_1 x_{11}, y_2 = x_{21}, y_3 = x_{32} + k_2 x_{31}, z = 2y_3 - y_1 y_2 - 2x_{2r} y_1 \tag{101.5}$$

$$\tau_1 = u_1 - u_{1r} + k_1 x_{12}, \tau_2 = u_2 - \dot{\overline{u}}_{2r}, v_2 = \dot{x}_2 - \overline{u}_{2r}, e_2 = y_2 - y_{2d} \tag{101.6}$$

and design the dynamic feedback control law as

$$\tau_1 = -k_4 y_1 - \tfrac{1}{2V} k_3 y_{2d} z, \tau_2 = \dot{v}_{2d} - (k_5 + k_6)\dot{e}_2 - k_5 k_6 e_2, \tag{101.7}$$

where $k_i (1 \le i \le 6)$ are the constants, y_{2d} is the assistant dynamic variable, $V = (y_1^2 + y_{2d}^2)/2$, then the closed-loop system is

$$\begin{aligned} &\dot{y}_1 = -k_4 y_1 - k_3 y_{2d} z / V \\ &\dot{y}_{2d} = -k_4 y_{2d} + k_3 y_1 z / V \\ &\ddot{e}_2 = \tau_2 - \dot{v}_{2d} = -(k_5 + k_6)\dot{e}_2 - k_5 k_6 e_2, \\ &\dot{x}_{31} = -k_2 x_{31} + z/2 + f \\ &\dot{z} = -2k_2^2 x_{31} - (k_3 - k_2) z - k_3 h z + f_4 \end{aligned} \tag{101.8}$$

where

$h = \frac{e_2 y_{2d}}{V}, f_2 = \frac{y_1 y_2}{2} + x_{2r} y_1, f_4 = f_3 - y_1 \dot{e}_2 - k_4 y_1 e_2, f_3 = f_1 + k_2 y_1 y_2 + 2k_2 x_{2r} y_1,$

$f_1 = 2(y_2 u_{1r} - k_1 x_2 x_{12} - y_1 \overline{u}_{2r}).$

The last two equations of (101.8) can be rewritten as

$$\dot{\boldsymbol{\eta}} = \boldsymbol{A}\boldsymbol{\eta} + \tilde{\boldsymbol{A}}(t)\boldsymbol{\eta} + \boldsymbol{f}(t), \tag{101.9}$$

where

$$\boldsymbol{\eta} = \begin{bmatrix} x_{31} \\ z \end{bmatrix}, \mathbf{A} = \begin{bmatrix} -k_2 & 1/2 \\ -2k_2^2 & -(k_3 - k_2) \end{bmatrix} \tilde{\mathbf{A}}(t) = \begin{bmatrix} 0 & 0 \\ 0 & -k_3 h \end{bmatrix}, \mathbf{f}(t) = \begin{bmatrix} f_2(t) \\ f_4(t) \end{bmatrix}. \tag{101.10}$$

101.3 Controller Design

101.3.1 Universal Exponential Regulator

Theorem 1 If (101.4) satisfies Hypothesis 1, $y_{2d}(0) \neq 0$, the parameters are chosen as

$$k_3 > k_2, \min(k_1, k_2, k_5, k_6) > k_4 > 0 \tag{101.11}$$

and the real parts of the eigenvalues of $\boldsymbol{A}$ are less than $-k_4$, then (101.7) guarantees

(1) $\mathbf{X}_e$ tends to be 0 exponentially, (2) $\dot{x}_2$ is bounded, and (3) (u_1, u_2) are bounded.

Proof By (101.8), there is $\dot{V} \triangleq \frac{\mathrm{d}V}{\mathrm{d}t} = -2k_1 V \Rightarrow V(t) = V(0)\mathrm{e}^{-2k_4 t}.V(0) \geq y_{2d}^2(0) > 0$; so $V(t) > 0$ for $t \in [0, \infty)$, and

$$\begin{aligned} &\max(|y_1(t)|, |y_{2d}(t)|) \leq \sqrt{2V(0)}\mathrm{e}^{-k_1 t} \Rightarrow (x_{11}, x_{12}) \in \mathrm{EXP}(k_4) \\ &(e_2, \dot{e}_2) \in \mathrm{EXP}(\lambda) \Rightarrow x_{21} = e_2 + y_{2d} \in \mathrm{EXP}(k_4), \lambda \triangleq \min(k_5, k_6). \end{aligned} \tag{101.12}$$

Due to Hypothesis 1, $f_i (1 \leq i \leq 4) \in EXP(k_4)$. Clearly, $h \in EXP(\lambda - k_4)$; thus, $\exists c_1 > 0 \Rightarrow \lim_{t \to \infty} \tilde{\mathbf{A}}(t) = 0, \int_0^\infty ||\tilde{\mathbf{A}}(t)|| \mathrm{d}t \leq c_1$. By Lemma 2, $\exists c_2 > 0$ so that

$$\max(|x_{31}(t)|, |z(t)|) \leq \|\boldsymbol{\eta}(t)\| \leq c_2 \mathrm{e}^{-k_4 t}. \tag{101.13}$$

Thus, (101.6) implies $y_3 \in EXP(k_4)$ and thereby $x_{32} \in EXP(k_4)$.

(2) As $\dot{e}_2$ converges exponentially and $\dot{x}_2 = \dot{e}_2 + \bar{u}_{2r}, \bar{u}_{2r}$ and $\dot{x}_2$ are bounded.
(3) Due to $(y_1, z, y_{2d}) \in EXP(k_4), V(t) \neq 0$, so (τ_1, τ_2) and (u_1, u_2) are bounded.
No PEC is needed in Theorem 1, so (101.7) is valid for point stabilization of (101.2).

101.3.2 Universal K-exponential Controller

Theorem 1 achieves no asymptotic stability. Denote $o \triangleq \left\| \bar{\boldsymbol{X}}_e(0) \right\|, \bar{\boldsymbol{X}}_e = [\boldsymbol{X}_e^{\mathrm{T}}, x_{22}]^{\mathrm{T}}$.

Theorem 2 If (101.4) satisfies Hypothesis 1, (101.11) is satisfied, the real parts of eigenvalues of $\mathbf{A}$ are less than $-k_4$, and

$$\dot{y}_{2d}(0) = y_{2d}(0) = \sqrt{-y_1^2(0) + \sqrt{y_1^4(0) + o^{4\varepsilon}}} \leq o^{\varepsilon}, 0 < \varepsilon < 1, \tag{101.14}$$

then the control law

$$\text{if } \bar{\mathbf{X}}_e(0) = 0, \text{ then } \quad u_1 = u_{1r}, u_2 = u_{2r}$$
$$\text{else} \begin{cases} \tau_1 = -k_4 y_1 - k_3 y_{2d} z / V \ \tau_2 = \dot{v}_{2d} - (k_5 + k_6)\dot{e}_2 - k_5 k_6 e_2 \\ v_{2d} \triangleq \dot{y}_{2d} = -k_4 y_{2d} + k_3 y_1 z / V \end{cases} \tag{101.15}$$

guarantees (1) $\mathbf{X}_e$ is K-exponentially stable, (2) $\dot{x}_2$ is bounded, and (3) (u_1, u_2) is bounded.

Proof It is easily known that only the case of $\bar{\mathbf{X}}_e(0) \neq 0$ needs discussion.

1. By (101.5) and $a^2 + b^2 \leq (a+b)^2 (ab \geq 0)$, then $| y_1(0) | \leqslant (1 + k_1)o$,

$$V(0) = \tfrac{1}{2}\sqrt{y_1^4(0) + o^{4\varepsilon}} \leq \tfrac{1}{2}\left(y_1^2(0) + o^{2\varepsilon}\right) \leq \tfrac{1}{2}\kappa_1^2(o) \tag{101.16}$$

where $\kappa_1(o) = (1 + k_1)o + o^{\varepsilon} \in \mathrm{FK}(o)$. By (101.8), there is $V(t) = V(0)\mathrm{e}^{-2k_4 t}$ and

$$\max(| y_1(t) |, | y_{2d}(t) |) \leq \sqrt{2V(t)} \leq \kappa_1(o)\mathrm{e}^{-k_4 t}. \tag{101.17}$$

By $\dot{x}_{11} = x_{12} = -k_1 x_{11} + y_1$ and Lemma 2, $\exists \kappa_2(o) \in FK(o) \Rightarrow | x_{11}(t) | \leqslant \kappa_2(o)\mathrm{e}^{-k_4 t}$, so

$$\left| x_{12}(t) \right| \leq \left| y_1(t) \right| + k_1 \left| x_{11}(t) \right| \leq \kappa_3(o)\mathrm{e}^{-k_4 t}, \tag{101.18}$$

where $\kappa_3(\cdot)=\kappa_1(\cdot)+k_1\kappa_2(\cdot)\in FK(o)$. By (101.8), there is

$\|[e_2(t),\dot{e}_2(t)]\|\le\kappa_4(\|[e_2(0),\dot{e}_2(0)]\|)\mathrm{e}^{-\lambda t}$,

where $\kappa_4(\cdot)\in \mathrm{FK}(o)$. By $e_2=y_2-y_{2d}$, $y_2=x_{21}$, there is

$$\begin{aligned}&|e_2(0)|\le|x_{21}(0)|+|y_{2d}(0)|\le o+o^{\varepsilon}\ |\dot{e}_2(0)|\le|x_{22}(0)|+|\dot{y}_{2d}(0)|\le o+o^{\varepsilon}\\&\Rightarrow\kappa_4\left(\|[e_2(0),\dot{e}_2(0)]\|\right)\le\kappa_4\left(|e_2(0)|+|\dot{e}_2(0)|\right)\le\kappa_5(o),\end{aligned}\tag{101.19}$$

where $\kappa_5(o)=\kappa_4(2o+2o^{\varepsilon})\in FK(o)$, so $\|[e_2(t),\dot{e}_2(t)]\|\le\kappa_5(o)\mathrm{e}^{-\lambda t}$. Thus

$$|y_2(t)|=|x_{21}(t)|\le|e_2(t)|+|y_{2d}(t)|\le\kappa_6(o)\mathrm{e}^{-k_4t},\tag{101.20}$$

where $\kappa_6(\cdot)=\kappa_5(\cdot)+\kappa_1(\cdot)\in FK(o)$. By definition of $f_i(1\le i\le 4)$ and Hypothesis 1, there is $\max(|f_i(t)|)\le\kappa_7(o)\mathrm{e}^{-k_4t}$, where $\kappa_7(o)\in FK(o)$.

By (101.9), as $\lim_{t\to\infty}\tilde{A}(t)=0,\int_0^{\infty}||\tilde{A}(t)||\mathrm{d}t\le c_3(c_3>0),||f(t)||\le 2\kappa_7(o)\mathrm{e}^{-k_4t}$, Lemma2impliesthatthereis$\kappa_8(\cdot)\in\mathrm{FK}(o)$suchthat$\max(|x_{31}(t)|,|z(t)|)\le\kappa_8(o)\mathrm{e}^{-k_4t}$. By (101.6), there is $|y_3(t)|\le\frac{1}{2}(|z(t)|+|y_1(t)y_2(t)|+2|x_{2r}(t)y_1(t)|)$. Thus, there exists $\kappa_9(\cdot)\in FK(o)$ so that $|y_3(t)|\le\kappa_9(o)\mathrm{e}^{-k_4t}$; therefore,

$$|x_{32}(t)|\le|y_3(t)|+k_2|x_{31}(t)|\le\kappa_{10}(o)\mathrm{e}^{-k_4t},\tag{101.21}$$

where $\kappa_{10}(\cdot)=\kappa_9(\cdot)+k_2\kappa_8(\cdot)\in FK(o)$. By definition of $\mathbf{X}_e$, there is

$$\|\mathbf{X}_e(t)\|\le|x_{11}(t)|+|x_{12}(t)|+|x_{21}(t)|+|x_{31}(t)|+|x_{32}(t)|\le\kappa(o)\mathrm{e}^{-k_4t},\tag{101.22}$$

where $\kappa(\cdot)=\kappa_2(\cdot)+\kappa_3(\cdot)+\kappa_6(\cdot)+\kappa_8(\cdot)+\kappa_{10}(\cdot)\in\mathrm{FK}(o)$.

2. As $\bar{u}_{2r}$ is bounded and $\dot{x}_2=\dot{e}_2+\bar{u}_{2r}$, so $\dot{x}_2$ is bounded.

3. By (101.5)–(101.6), there is $\lim_{o\to0}y_1=\lim_{o\to0}x_{12}=0$,

$$\lim_{o\to0}|z|\le[2(1+k_2)+(1+k_1)(2|x_{2r}(0)|+o)]o.\tag{101.23}$$

By(101.14)and(101.16),thereis$\lim_{o\to0}|\frac{y_{2d}(0)z(0)}{V(0)}|\le\lim_{o\to0}\frac{o^{\varepsilon}[2(1+k_2)+(1+k_1)(2|x_{2r}(0)|+o)]o}{0.5o^{2\varepsilon}}=0$.

By (101.15), if $o\to0$, then $\tau_1(0)\to0$. Using (101.6), then $u_1(0)\to u_{1r}(0)$. Thus, u_1 is continuous and so is u_2. By (101.16)–(101.17), $\lim_{t\to\infty}|\frac{y_{2d}(t)z(t)}{V(t)}|\le\lim_{t\to\infty}\frac{\kappa_1(o)\kappa_8(o)}{0.5o^{2\varepsilon}}$

. When $o \neq 0, \lim_{t\to\infty} \frac{\kappa_1(o)\kappa_8(o)}{0.5o^{2\varepsilon}}$ is a bounded positive number. Thus, τ_1 is bounded according to (101.15) and so is τ_2. So (u_1, u_2) are bounded by (101.6).

In Theorem 2, $V(t)$ is strictly monotone convergent but $(y_{2d}(t), z(t))$ are not, as they may cross zero many times. It may happen that $V(t)$ decreases while $|\tau_1(t)|$ increases, which causes peak phenomena and oscillation.

101.3.3 Practical Universal Controller

In order to avoid oscillation, the control objective is loosened to practical stability; thus, there is no need to configure $y_{2d}(0)$ (only $y_{2d}(0) \neq 0$ is needed).

Theorem 3 If (101.4) satisfies Hypothesis 1, (101.11) is satisfied, $y_{2d}(0) \neq 0$, and the real parts of the eigenvalues of $\mathbf{A}$ are less than $-k_4$, then the smooth control law

$$\begin{cases} \tau_1 = -k_4 y_1 - k_3 y_{2d}(z + f_5)/(V + \delta) \\ \tau_2 = \dot{v}_{2d} - (k_5 + k_6)\dot{e}_2 - k_5 k_6 e_2, \qquad \delta > 0 \, f_5 = (2k_2 f_2 - 2k_2^2 x_{31} + f_1)/k_3 \\ v_{2d} \triangleq \dot{y}_{2d} = -k_4 y_{2d} + k_3 y_1 (z + f_5)/(V + \delta) \end{cases} \tag{101.24}$$

guarantees (1) $\mathbf{X}_e(t)$ is globally exponentially convergent, (2) $\dot{x}_2(t)$ is bounded, (3) (u_1, u_2) are bounded, and (4) $\lim_{t\to\infty}(u_1(t), u_2(t)) = (u_{1r}(t), u_{2r}(t))$.

Proof The proof is similar to that in Theorem 1 and 2. Notice that existence of δ eliminates oscillation and peaking phenomena in Theorem 2.

101.4 Simulation

Simulation is carried out on (101.2) and Theorem 3. The initial states are

$$(x_1, x_2, x_3, v_1, v_2, v_3) = (1,1,1,0,0,0), (x_{1r}, x_{2r}, x_{3r}, v_{1r}, v_{2r}, v_{3r}) = (1,1,1,0,0,0), \tag{101.25}$$

where $(v_i, v_{ir})(i = 1,2,3)$ denote $(\dot{x}_i, \dot{x}_{ir})$. The error limit is $\delta = 0.1$. The parameters are selected as $(k_1, k_2, k_3, k_4, k_5, k_6, k_7) = (2, 12/7, 7, 0.5, 2.2, 2.4)$.

1. Practical Trajectory Tracking

The reference inputs are $u_{1r} = \sin(\pi t/2), u_{2r} = \cos(\pi t/2)$ and the simulation results are shown in Fig. 101.1. It shows that the tracking errors converge to 0 exponentially except that x_{22} is bounded; besides, (u_1, u_2) converge to (u_{1r}, u_{2r}) exponentially. The peaking phenomena are eliminated completely.

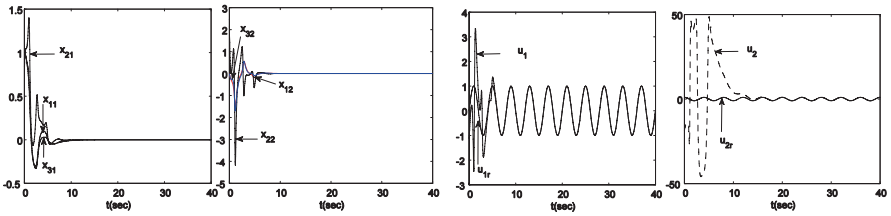


Fig. 101.1 Simulation results of practical trajectory tracking

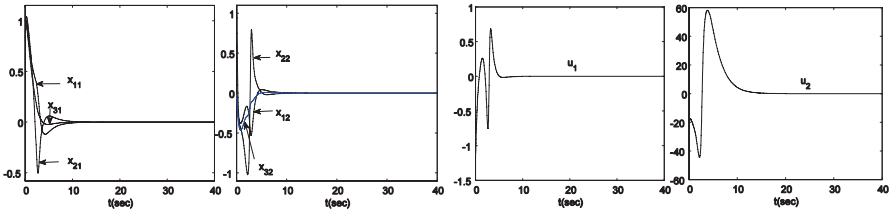


Fig. 101.2 Simulation results of practical point stabilization

2. Practical Point Stabilization

The desired point is the origin and the simulation results are shown in Fig. 101.2. It shows that all the state errors converge to 0 exponentially except that x_{22} is bounded; besides, the inputs (u_1, u_2) converge to 0 exponentially.

101.5 Conclusion

Three universal controllers are given for a second-order NS to get exponential convergence, asymptotic exponential stability, and practical stability, respectively. They can be applied to the point stabilization and the trajectory tracking without any modification. Theoretically, the asymptotic stability is better than the practical stability. Oscillation and peaking phenomena in Theorem 2 show that it is difficult to achieve the asymptotic stability and that practical stability may be the better choice.

As to the high-dimension NS, it is difficult to find a proper Lyapunov function and further research is demanded to generalize the method described in this chapter.

References

1. Do KD. Practical control of underactuated ships. Ocean Eng. 2010;37(13):1111–9.
2. Brockett RW. Asymptotic stability and feedback stabilization. In: Brockett RW, Millman RS, Sussmann HH, editor. Differential geometric control theory. Boston: Birkhauser; 1983. pp. 181–91.

3. Thuilot B, d'Andrea-Novel B, Micaelli A. Modeling and feedback control of mobile robots equipped with several steering wheels. IEEE Trans Rob Autom. 1996;12(3):375–90.
4. Kim DH, Oh JH. Tracking control of a two-wheeled mobile robot using input-output linearization. Control Eng Pract. 1999;7(3):369–73.
5. Kanayama Y, Nilipour A, Lelm CA. A locomotion control method for autonomous vehicles. Proceedings of IEEE International Conference on Robotics and Automation; New York: IEEE; 1988. pp. 1315–7.
6. Morin P, Samson C. Stabilization of trajectories for systems on Lie groups: application to the rolling sphere. Proceedings of the 17th world congress on the international federation of automatic control; London: Pergamon Press; 2008. pp. 508–13.
7. Lee TC, Song KT, Lee HC, et al. Tracking control of unicycle-modeled mobile robots using a saturation feedback controller. IEEE Trans Control Syst Technol. 2001;9(2):305–18.
8. Do KD, Jiang ZP, Pan J. Simultaneous tracking and stabilization of mobile robots: an adaptive approach. IEEE Trans Automat Contr. 2004;49(7):1147–52.
9. Dixon WE, Darken DM, Zhang FM, et al. Global exponential tracking control of a mobile robot system via a PE condition. IEEE Trans Syst Man Cybern B Cybern. 2000;30(1):120–42.
10. Bloch A, Drakunov S. Stabilization and tracking in the nonholonomic integrator via sliding modes. Syst Control Lett. 1996;29(2):91–9.
11. Do KD, Jiang ZP, Pan J. Universal controllers for stabilization and tracking of underactuated ships. Syst Control Lett. 2002;47(4):299–317.
12. Do KD, Jiang ZP, Pan J. A global output-feedback controller for simultaneous tracking and stabilization of unicycle-type mobile robots. IEEE Trans Rob Autom. 2004;20(3):589–94.
13. Tian YP, Cao KC. Time-varying linear controllers for exponential tracking of non-holonomic systems in chained form. Int J Robust Nonlinear Control. 2006;17(7):631–47.
14. Cao KC. Global K-exponential tracking control of nonholonomic systems in chained-form by output feedback. Acta Automatica Sinica. 2009;35(5):568–76.
15. Oriolo G, De Luca A, Vendittelli M. WMR control via dynamic feedback linearization: design, implementation, and experimental validation. IEEE Trans Control Syst Technol. 2002;10(6):835–52.
16. Yang EF, Gu DB, Mita T, Hu HS. Nonlinear tracking control of a car-like mobile robot via dynamic feedback linearization. Control 2004, UK: University of Path; 2004. ID–218.
17. Morin P, Samson C. Practical stabilization of driftless systems on Lie groups: the transverse function approach. IEEE Trans Autom Control. 2003;48(9):1496–508.
18. Li MJ. Exponential stabilization of nonholonomic systems [D]. Beijing: Beihang University; 2011 (In Chinese).

Chapter 102
Passivity and Biological Control of the Singular Mammal Body Complexity System

Peiyong Liu, Li Yang and Chao Liu

Abstract It is very difficult to use exact mathematical models to study complexity systems. As the units beside complex systems can interact to bring forth its complexity, it is completely true for the mammal body, a classic kind of biological complex system. As singular biological complexity systems, including the mammal body, are considered in a series of new research fields, a few passivity or biological control studies have been carried out up to now. In this research, the single-chamber model of the environmental hormone formaldehyde which is flowing in the mammal body has been set up according to the corresponding physiological rules with the model passivity described in detail. Under the strict passivity station, the feedback controller for this singular mammal body complexity system has been designed with a controller example also given as a model instantiation. Both passivity study and feedback controller design of the mammal body complex system can be applied to biological complexity systems and lay a useful foundation for singular system research.

Keywords Passivity · Biological control · Singularity system · Complexity system

102.1 Introduction

In a complexity system, there will be very complicated interaction; but nearly all the complex systems are constituted by simple components. Thus, complex systems are considered complicated, but complete [1]. Many fields of practical systems are

P. Liu (✉)
College of Life and Health Sciences, Northeastern University, 110819 Shenyang, China
e-mail: liupy@mail.neu.edu.cn

L. Yang
Department of Mathematics, Liaoning University, 110000 Shenyang, China
e-mail: yangli2923@163.com

C. Liu
Institute of Systems Science, Northeastern University, 110000 Shenyang, China
e-mail: liuchao@mail.neu.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_102

included in complex systems such as social system, financial system, ecosystem, and biosystem, including the mammal body [2].

Endocrine disruptors or environmental hormones are a kind of chemicals that disrupt and influence most of the systems in biological complex systems. The interacting impact of systems may mainly involve neural system, immune system, and reproductive system of the mammal body. Thus, these environmental hormones may affect the survival problems of animals, especially of mammals including human beings [3].

Formaldehyde is very common in our daily life. It is extensively adopted by many industries, especially paint factories; thus, the atmosphere in some new buildings and decorated environments are full of this substance. After absorbing formaldehyde, persons may feel dizziness, sicchasia, and feebleness. In this sense, this model is very useful for studying formaldehyde flowing in the mammal body and also to carry out intensive research of the essential rules of formaldehyde flowing in the mammal body; it is also important for later quantitative studies of biological complex systems. Up to now, few quantitative models have been set up to explain complex systems, including the mammal body [4–10].

Singularity systems are a series of dynamical systems featuring differential algebraic. In fact, economics, electronic power, and space navigation are all singularity system models [11]. Passivity study and passivity of singularity systems have been researched effectively. But applications of passivity theorems for singularity systems, especially biological complex systems, have been dealt with little [12–19]. The feedback controller and the model passivity of this chapter can be used to lay the theoretical foundation for solving the environmental hormone pollution problems of formaldehyde that remained in mammals, especially in the human body.

102.2 Descriptive Poly-Chamber Model of the Mammal Body Complex System

A multi-chamber model was established on the basis of physical processes of the environmental hormone formaldehyde flowing in the mammal body as shown in Fig. 102.1.

The notes in Fig. 102.1 can be considered as follows:

1. The arrows $\Rightarrow$ and $\Leftrightarrow$ in Fig. 102.1 refer to unidirectional and bidirectional flowing of formaldehyde in the mammal body, respectively.
2. The Arabic numerals and letters in Fig. 102.1 refer to the systems or tissues affected by hydroxyl formaldehyde. 1CS refers to the circulating system; 2TI to the tissues in the human body; 3RS to the systema respiratorium; IN is the input; and 0 is the environment.
3. Blood is the central cavity that can carry formaldehyde around the organs. The input of formaldehyde is from the systema respiratorium mainly.
4. Some parts of formaldehyde are rereleased by respiratory activities (y_1), while some parts form adducts with protein, nucleic acids, and other molecules directly. The formaldehyde that does not react with molecules can be oxidized quickly to

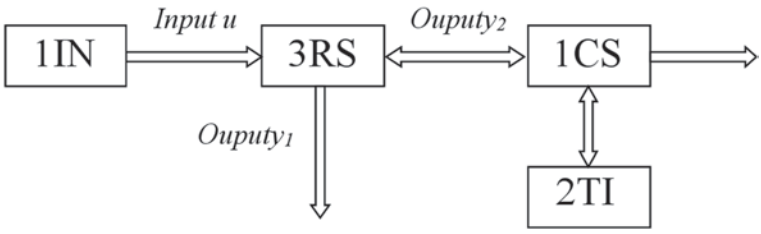


Fig. 102.1 Multi-chamber conceptual model of formaldehyde flowing route in mammal body

formate by two different enzyme systems. And formic acid (or formate) is further oxidized to CO_2 that can be breathed out from the body, or can be eliminated in the form of sodium formate with urine from the body (y_2). And small parts of formaldehyde still remain in the body to be distributed to various tissues.

Let x_1, x_2, and x_3 refer to the amount of formaldehyde in different organs; θ_{ij} refers to the flowing rate constant values from chamber i to chamber $j(i, j = 1, 2, 3)$; "In" refers to "input"; dynamics of first class can be used to describe the flowing process of formaldehyde [14]. According to the relationships among the chambers in Fig. 102.1, equations of the state can be derived as follows:

$$\begin{cases} \dot{x}_1(t) = -(\theta_{21} + \theta_{10})x_1 + \theta_{21}x_2 + \theta_{31}x_3 \\ \dot{x}_2(t) = \theta_{12}x_1 - \theta_{21}x_2 \\ \dot{x}_3(t) = In - (\theta_{31} + \theta_{30})x_3 \end{cases}. \tag{102.1}$$

And the corresponding output equation relative to Eq. (102.1) can be written as:

$$\begin{cases} y_1(t) = \theta_{30}x_3 \\ y_2(t) = \theta_{10}x_1 \end{cases}. \tag{102.2}$$

Equation (102.1) can be changed to (102.3).

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases}, \tag{102.3}$$

where $A = \begin{bmatrix} -\theta_{12} - \theta_{10} & \theta_{21} & \theta_{31} \\ \theta_{12} & -\theta_{21} & 0 \\ 0 & 0 & -\theta_{31} - \theta_{30} \end{bmatrix}; B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}; C = \begin{bmatrix} 0 & 0 & \theta_{30} \\ \theta_{10} & 0 & 0 \end{bmatrix}; D = 0.$

Let $u(t) = In(t)$. Another simple equation can be used as below:

$$x_4 = In = x_1 + x_2 + x_3 + y_1 + y_2. \tag{102.4}$$

If the number of dimensions of the output and input is the same, using singularity system theory, the model (102.3) can be written as a singular model as below:

$$\begin{cases} \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{B}_1\mathbf{u}(t) + \mathbf{B}_2\mathbf{w}(t) \\ \mathbf{y}(t) = C_1\mathbf{x}(t) + \mathbf{D}_1\mathbf{u}(t) + \mathbf{D}_2\mathbf{w}(t) \end{cases}, \tag{102.5}$$

where $\mathbf{x}(t)$, $\mathbf{u}(t)$, and $\mathbf{w}(t)$ represent states, control input, and ectogenic input for the common singularity system (102.5), respectively.

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{102.6}$$

A_1, B_1, and C_1 can be calculated using A, B. If it shows:

$$A_1 = \begin{bmatrix} -\theta_{12} - \theta_{10} & \theta_{21} & \theta_{31} & 0 \\ \theta_{12} & -\theta_{21} & 0 & 0 \\ 0 & 0 & -\theta_{31} - \theta_{30} & 0 \\ 1 + \theta_{10} & 1 & 1 + \theta_{30} & -1 \end{bmatrix}, \tag{102.7}$$

and

$$B_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \; C_1 = \begin{bmatrix} 0 & 0 & \theta_{30} & 0 \\ \theta_{10} & 0 & 0 & 0 \end{bmatrix}, \tag{102.8}$$

then B_2, D_1, and D_2 can be obtained as follows:

$$B_2 = 0, \; D_1 = 0, \; D_2 = 0.3. \tag{102.9}$$

102.3 Passivity Study and Feedback Controller Design in Singular Mammal Complexity Systems

A definition of the system (102.5) can be given as follows:

Definition 1 [4] If the two functions $V(\mathbf{x})$ and $W(\mathbf{x})$ exist and can be positive semi-definite and positive definite separately, then they contain

$$\int_0^\tau \mathbf{w}^{\mathrm{T}}\mathbf{y}dt \geq V(x(\tau)) - V(\mathbf{x}(0)) + \int_0^\tau W(\mathbf{x})dt. \tag{102.10}$$

The state space will be strictly passive, under the condition of any initial value $\mathbf{x}(0)$ or time τ.

System (102.1) can be differentiated from any side about τ to determine the passivity. And if

$$\mathbf{w}^{\mathrm{T}}\mathbf{y} \geq \dot{V}(\mathbf{x}) + W(\mathbf{x}), \tag{102.11}$$

and if the state space is strictly passive, (102.11) must be correct for all initial values $\mathbf{x}(0)$ and times τ.

After calculation, it is testified that the Lyapunov function of the system (102.1) can be nonnegative. We can let $V = \mathbf{x}^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\mathbf{P}\mathbf{x}$, where $\mathbf{P}$ is to be determined and satisfy $\mathbf{E}^{\mathrm{T}}\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{E} \geq \mathbf{0}$.

Theorem1 As to one free system ($\mathbf{u}(t) \equiv \mathbf{0}$), if it satisfies $\mathbf{D}_2 + \mathbf{D}_2^{\mathrm{T}} > \mathbf{0}$, $\mathbf{E}^{\mathrm{T}}\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{E} \geq \mathbf{0}$, then there is a feasible solution for the linear matrix inequality

$$\begin{pmatrix} A_1{}^{\mathrm{T}}\mathbf{P} + \mathbf{P}^{\mathrm{T}}A_1 + \delta\mathbf{I} & \mathbf{P}^{\mathrm{T}}\mathbf{B}_2 - \frac{1}{2}C_1{}^{\mathrm{T}} \\ \mathbf{B}_2^{\mathrm{T}}\mathbf{P} - \frac{1}{2}C_1 & -\frac{1}{2}\mathbf{D}_2 - \frac{1}{2}\mathbf{D}_2^{\mathrm{T}} \end{pmatrix} \leq \mathbf{0}. \tag{102.12}$$

Where $\delta > 0$, the singular system will be straitly passive, and the system (102.5) can be considered as a closed-loop system.

Theorem2 As to one singular system, when $\delta > 0$, $\mathbf{D}_2 + \mathbf{D}_2^{\mathrm{T}} > \mathbf{0}$, $\mathbf{E}^{\mathrm{T}}\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{E} \geq \mathbf{0}$, an appropriate solution can be used to solve the problem of a linear matrix inequality

$$\begin{pmatrix} (A_1\mathbf{X} + \mathbf{B}_1\mathbf{W})^{\mathrm{T}} + (A_1\mathbf{X} + \mathbf{B}_1\mathbf{W}) & \mathbf{B}_2 - \frac{1}{2}(C_1\mathbf{X} + \mathbf{D}_1\mathbf{W})^{\mathrm{T}} & \mathbf{X}^{\mathrm{T}} \\ \mathbf{B}_2^{\mathrm{T}} - \frac{1}{2}(C_1\mathbf{X} + \mathbf{D}_1\mathbf{W}) & -\frac{1}{2}\mathbf{D}_2 - \frac{1}{2}\mathbf{D}_2^{\mathrm{T}} & \mathbf{0} \\ \mathbf{X} & \mathbf{0} & -\delta^{-1}\mathbf{I} \end{pmatrix} < \mathbf{0}. \tag{102.13}$$

And the closed-loop system should be straitly passive. If $\mathbf{W}^*$, $\mathbf{X}^*$ is one of the proper solutions of (102.13), the feedback controller $\mathbf{u} = \mathbf{K}\mathbf{x} = \mathbf{W}^*(\mathbf{X}^*)^{-1}\mathbf{x}$ can be used to show the strict passivity.

102.4 Feedback Controller Example of the Passivity State in Poly-Chamber Models

The feedback controller designed above can be applied in the singular system (102.5). On the basis of the physiological process laws of formaldehyde flowing in the mammal body, θ_{ij} in A_1 and C_1 are set as below, and let $\theta_{10} = 0.05$, $\theta_{12} = 0.03$, $\theta_{21} = 0.001$, $\theta_{30} = 0.51$, and $\theta_{31} = 0.36$.

$$A_1 = \begin{bmatrix} -0.08 & 0.001 & 0.36 & 0 \\ 0.03 & -0.001 & 0 & 0 \\ 0 & 0 & -0.87 & 0 \\ 1.05 & 1 & 1.51 & -1 \end{bmatrix}, C_1 = \begin{bmatrix} 0 & 0 & 0.51 & 0 \\ 0.05 & 0 & 0 & 0 \end{bmatrix} \quad (102.14)$$

And

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, B_2 = 0, D_1 = 0, D_2 = 0.1 \quad (102.15)$$

Through calculations, on the basis of theorem 2, we can easily get $\delta = 0.0468$ and

$$W = \begin{bmatrix} -0.4806 & 0.0071 & -6.7960 & -3.9630 \end{bmatrix}. \quad (102.16)$$

So, we can describe the strictly passive state feedback controller by

$$u = Kx = \begin{bmatrix} -1.1820 & -1.4957 & -2.4637 & -0.4852 \end{bmatrix} x. \quad (102.17)$$

102.5 Conclusion

In this study, a single-chamber model of environmental hormone formaldehyde flowing in the mammal body has been set up according to the corresponding physiological rules with the model passivity described in detail. Under the strict passivity station, the feedback controller of the singular mammal body complexity system has been designed with a controller example also given as model instantiation.

The mammal body complexity system is also described using descriptor system theory. It can be applied to biological complex systems and lay useful a foundation for singular system research; furthermore, the feedback controller in this chapter can also be applied to find proper tools for solving environmental hormone pollution problems such as formaldehyde, benzene, etc. that remain in the human body.

References

1. Mitchell W. Complexity: the emerging science at the edge of order and chaos. New York: Tocuhstone; 1992. p. 20–58.
2. Zhang SY. A brief introduction to complex systems and complexity science. J Qingdao Univ. 2001;16(4):25–8. (In Chinese).

3. Diamanti-Kandarakis E, Bourguignon JP, Giudice LC. Endocrine-disrupting chemicals: an Endocrine Society scientific statement. Endocr Rev. 2009;30(4):293–342.
4. Harrison PT. Reproductive health in humans and wildlife: are adverse trends associated with environmental chemical exposure. Sci Total Environ. 1997;205(1):97–106.
5. Jimenez B. Environmental effects of endocrine disruptors current methodologies for assessing wildlife health effects. Trends Anal Chem. 1997;16(3):596–606.
6. Liu PY, Zhang QL, Zhang Y. Analysis and optimization on poly-chambers models of complex control system. Control Eng China. 2005;12(2):109–11. (In Chinese).
7. Li GF, Li HY. Observer-based passive control for uncertain linear systems with delay in state and control input. Chin Phys. 2005;14(15):2379–86. (In Chinese).
8. Rgensen SEJ. Fundamentals of ecological modeling. Shanghai: Translation Press Company; 1988. p. 34–50.
9. Tsuguchika K. Modeling of signaling pathways for endocrine disruptors. BioSystems. 2000;5(1):23–31.
10. Warren GF. Human development exposure to endocrine active compounds. Environ Toxicol Pharmacol. 2002;12(1):75–81.
11. Belura MN, Trentelmanb HL. The strict dissipativity synthesis problem and the rank of the coupling QDF. Syst Control Lett. 2004;51(1):247–58.
12. Dong XZ, Zhang QL. Robust passive control for singular systems with time-varying uncertainties. Control Theory and Appl. 2004;21(3):517–20.
13. Dong XZ, Zhang QL, Zhao LC. Passive control of linear singular systems. J Biomathematics. 2004;19(1):185–7. (In Chinese).
14. Fridman E, Shaked U. On delay-dependent passivity. IEEE Trans Automat Control. 2002;47(2):664–9.
15. Yang DM, Zhang QL. Singular system. Beijing: Science Press; 2004. p. 25–36. (In Chinese).
16. Werner K. Amphibians as a model to study endocrine disruptors: âestrogenic activity of environmental chemicals in vitro and in vivo. Sci Total Environ. 1999;225(1):59–68.
17. Xie SL, Xie H. Robust dissipative control for linear systems with dissipative uncertainty and nonlinear perturbation. Syst Control Lett. 1997;29(2):255–68.
18. Yu L, Pan HT. Robust passive control of linear systems with time-varying uncertain parameters. Acta Automatica Sinca. 1998;24(9):368–72. (In Chinese).
19. Yang DM, Zhang QL. Passive control of linear singular systems via output feedback. J Northeast Univ (Nat Sci). 2004;25(2):310–3. (In Chinese).

Chapter 103
$H\infty$ Control for NRPCS Based on the Takagi-Sugeno Fuzzy Model

Cheng Gong

Abstract In this chapter, we use the Takagi-Sugeno (T-S) fuzzy model to model the nuclear reactor power control system (NRPCS), which is nonlinear time-varying and not easy to control. First, we give the point-kinetic nonlinear time-varying model of the NRPCS; then we choose the reactor power as the premise variable, propose the membership, and present a T-S fuzzy model for the NRPCS. Finally, an H∞ controller is investigated. The numerical example illustrates the advantage of the proposed model.

Keywords T-S fuzzy model · H∞ · Reactor power control

103.1 Introduction

For the safety a of nuclear power plant (NPP), the nuclear reactor power control system (NRPCS) plays an important role. Many control methods have been investigated in the area of nuclear reactor control. For example, some authors proposed the controller design method for an NPP with 15 % variation of reactor power about nominal power (100 %) [1]. Some authors used the algorithm of predictive control in the NRPCSs [2]. And the fuzzy proportional-integral-derivative (PID) control method for the NRPCSs was studied [3, 4]. The above-mentioned methods are all conventional control.

In contrast, some authors extended the modern control theorem to the NRPCSs. Some scholars demonstrated the improved robustness of state feedback adaptive control (SFAC) to handle changes of reactor parameters [5]. An adaptive optimal controller design for a NRPCS of a pressurized water reactor (PWR) was proposed [6]. Many authors researched the method of SFAC for NRPCSs [7, 8].

As we all know, the NRPCSs are in nature nonlinear, with their parameters varying with time as a function of power level, fuel loss, and control rod worth. The

C. Gong (✉)
School of Mathematical Science, Heilongjiang University, 150080 Harbin, China
e-mail: gongcheng2004@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_103

authors took account of power variations in the power plant and presented a neural network controller which displayed good performance and stability in a wide range of operations [9].

At the same time, the model is very effective to obtain nonlinear control systems. Many applications on T-S control were studied in many areas. For example, consumer products and industrial processes [10, 11]. In view of these facts, we will model the NRPCSs with the T-S fuzzy model which can be handled with the modern control theory frame. Based on Lyapunov's functional approach and the T-S fuzzy model, we present the H∞ controller for the NRPCSs with all results proposed in terms of linear matrix inequality (LMIs).

103.2 System Descriptions and Preliminaries

In this section, we use the model on the nominal working point n_r and consider the disturbance in the coolant temperature reactivity. The model is shown as follows [7]:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1 u(t) + B_2 \omega(t), \\ y(t) &= Cx(t) \end{aligned} \tag{103.1}$$

$$A = \begin{bmatrix} -\frac{\beta}{\lambda} & \frac{\beta}{\lambda} & \frac{n_r \alpha_f}{\Lambda} & \frac{n_r \alpha_c}{2\Lambda} \\ \lambda & -\lambda & 0 & 0 \\ \frac{f_f P_{0a}}{\mu_f} & 0 & \frac{-\Omega}{\mu_f} & \frac{\Omega}{2\mu_f} \\ \frac{(1-f_f)P_{0a}}{\mu_c} & 0 & \frac{\Omega}{\mu_c} & \frac{-(2M+\Omega)}{2\mu_c} \end{bmatrix} \quad B_1 = \begin{bmatrix} n_r G_r & 0 & 0 & 0 \end{bmatrix}^T$$

where $B_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$

$C = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$

$x = \begin{bmatrix} \Delta n_r & \Delta C_r & \Delta T_f & \Delta T_l \end{bmatrix}^T$

$u = [z_r]$.

The variables are shown as below:

n_r is the neutron density near to the equilibrium point, Δn_r is the small fluctuation of n_r, and C_r is the precursor density near to initial equilibrium density.

ΔC_r is the small fluctuation of C_r. T_f is the reactor fuel temperature, ΔT_f is the small fluctuation of T_f. T_l is the temperature of the water and ΔT_l is the small fluctuation of T_l. z_r is the control rod speed (fraction of core length per second). G_r is the reactivity worth of the rod unit length. $u(t)$ is the control input. $\omega(t)$ is the disturbance of coolant temperature.

The parameters represent the Three Mile Island-type reactor [9].

$$\begin{aligned}&\beta = 0.006019,\ \lambda = 0.15s^{-1},\ \Lambda = 2\times10^{-5}s,\ f_f = 0.92\\&P_{0a} = 2500MW,\ \mu_f = 26.3MWs/^0C,\ G_r = 0.0145\end{aligned} \tag{103.2}$$

$$\begin{aligned}&\alpha_f = (n-4.24)\times10^{-5} = -3.24\times10^{-5},\\&\alpha_c = (-4n_r-17.3)\times10^{-5},\ \Omega = (\frac{5}{3}n_r-4.933)MWs/°C,\\&M = (28n_r+74)MWs/°C,\ \mu_c = (\frac{160}{9}n_r-54.022)MWs/°C\end{aligned} \tag{103.3}$$

Here, we want to model System (103.1) with the T-S Fuzzy model which can be handled with the modern control theory frame. Then, we design a common control law $u(t)$ for System (103.1) with H∞ sense.

103.3 Main Results

103.3.1 Modeling

As to System (103.1), we can know that the equation varies with the working point n_r; so, we choose some functions of the working point n_r as the premise variable ($z_1(t) \sim z_7(t)$). Use the method for construction of the fuzzy model [12].

Define

$$z_1(t) \equiv \frac{n_r\alpha_f}{\Lambda}, z_2(t) \equiv \frac{n_r\alpha_c}{2\Lambda}, z_3(t) \equiv \frac{-\Omega}{\mu_f}, z_4(t) \equiv \frac{(1-f_f)P_{0a}}{\mu_c}, \tag{103.4}$$

and

$$z_5(t) \equiv \frac{\Omega}{\mu_c}, z_6(t) \equiv \frac{-(2M+\Omega)}{2\mu_c}, z_7(t) \equiv n_rG_r.$$

Then, we have

$$\begin{aligned}\dot{x}(t) &= \begin{bmatrix}-\beta/\lambda & \beta/\lambda & z_1(t) & z_2(t)\\ \lambda & -\lambda & 0 & 0\\ f_fP_{0a}/\mu_f & 0 & z_3(t) & -z_3(t)/2\\ z_4(t) & 0 & z_5(t) & z_6(t)\end{bmatrix}x(t) + \begin{bmatrix}z_7(t)\\0\\0\\0\end{bmatrix}u + \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\omega(t).\\ y &= \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}x(t)\end{aligned} \tag{103.5}$$

Next, calculate the minimum and maximum values of $z_1(t) \sim z_7(t)$ under $n_r \in [15\% \quad 105\%]$ (we suppose that the power variation is from 15 to 105 %). They are obtained as follows: $\max_{n_r\in[0.15,1.05]} z_1(t) = -0.3068$ $\min_{n_r\in[0.15,1.05]} z_1(t) = -1.6748$

From the maximum and minimum values, $z_1(t)$ can be represented as: $z_1(t) = M_{big}(z_1(t))*(-0.3068) + M_{little}(z_1(t))*(-1.6748)$,

where $M_{big}(z_1(t)) + M_{little}(z_1(t)) = 1$.

Therefore, the membership function can be calculated as $M_{big}(z_1(t)) = \dfrac{z_1(t)+1.6748}{1.368}, M_{little}(z_1(t)) = -\dfrac{z_1(t)+0.3068}{1.368}$.

We take the same procedure for $z_2(t) - z_7(t)$ as well. We obtain the general form of the membership function

$$M_{big}(z_i(t)) = \frac{z_i(t) - \min(z_i(t))}{\max(z_i(t)) - \min(z_i(t))}, M_{little}(z_i(t)) = \frac{\max(z_i(t)) - z_i(t)}{\max(z_i(t)) - \min(z_i(t))}, i = 2,\ldots 7 .$$

We name the membership functions "*big*" and "*little;*" then System (103.5) is represented by the fuzzy models.

Rule r ***If*** $z_1(t)$ is big, $z_2(t)$ is big, $z_3(t)$ is big, $z_4(t)$ is big, $z_5(t)$ is big, $z_6(t)$ is big, and $z_7(t)$is big, ***then*** $\dot{x}(t) = A_r x(t) + B_{1r} u + B_{2r}\omega(t)$.

$$y = C_r x(t)$$

In theory, there are 2^7 rules, i.e., 2^7 fuzzy sub-models. In the meanwhile, it should be pointed out that all the premise variables $z_i(t)$ vary according to $n_r \in [15\% \quad 105\%]$ monotonously. Based on this fact, we can reduce the rule number to two which are: Rule 1: n_r is about 15%. Rule 2: n_r is about 105%.

According to the above rules, we can get the T-S fuzzy model and rules about $z_i(t)$:

Rule 1 ***If*** $z_1(t)$ is big, $z_2(t)$ is big, $z_3(t)$ is big, $z_4(t)$ is big, $z_5(t)$ is little, $z_6(t)$ is big, and $z_7(t)$ is little, ***then*** $\dot{x}(t) = A_1 x(t) + B_{11} u + B_{21}\omega(t)$.

$$y = C_1 x(t)$$

Rule 2 ***If*** $z_1(t)$ is little, $z_2(t)$ is little, $z_3(t)$ is little, $z_4(t)$ is little, $z_5(t)$ is big, $z_6(t)$ is little, and $z_7(t)$ is big, ***then*** $\dot{x}(t) = A_2 x(t) + B_{12} u + B_{22}\omega(t)$.

$$y = C_2 x(t)$$

The following global T-S fuzzy model can be obtained:

$$\begin{cases} \dot{x}(t) = \dfrac{\sum_{i=1}^{2} \alpha_i(z(t))[A_i x(t) + B_{1i} u(t) + B_{2i}\omega(t)]}{\sum_{i=1}^{r} \alpha_i(z(t))} \\ \quad = \sum_{i=1}^{2} \mu_i(z(t))[Ax(t) + B_{1i} u(t) + B_{2i}\omega(t)] \\ y(t) = \sum_{i=1}^{2} \mu_i(z(t)) C_i x(t) \\ \quad = C(t)x(t), \end{cases}$$

where $\alpha_i(z(t)) = \prod_{j=1}^{7} F_{ij}(z_j(t))\mu_i(z(t)) = \alpha_i(z(t)) / \sum_{i=1}^{2} \alpha_i(z(t))F_{ij}(z_j(t))$, is the grade of membership of $z_j(t)$ in F_{ij}, and it is assumed that $\alpha_i(z(t)) \geq 0,\ i = 1,2, \sum_{i=1}^{2} \alpha_i(z(t)) > 0$ for all t.

Therefore, $\mu_i(z(t)) \geq 0$ and $\sum_{i=1}^{2} \mu_i(z(t)) = 1$ for all

$$t \cdot M_{big}(z_i(t)) = \frac{z_i(t) - \min(z_i(t))}{\max(z_i(t)) - \min(z_i(t))}, M_{little}(z_i(t)) = \frac{\max(z_i(t)) - z_i(t)}{\max(z_i(t)) - \min(z_i(t))}, i = 2,\ldots 7,$$

where

$$M_{big}(z_1(t)) = \frac{z_1(t) + 1.6748}{1.368}, M_{little}(z_1(t)) = -\frac{0.3068 + z_1(t)}{1.368},$$

$$M_{big}(z_3(t)) = \frac{z_3(t) + 0.2541}{0.0570}$$

$$M_{little}(z_3(t)) = -\frac{0.1971 + z_3(t)}{0.0570}, M_{big}(z_4(t)) = \frac{z_4(t) - 2.7515}{0.7766},$$

$$M_{little}(z_4(t)) = \frac{3.5280 - z_4(t)}{0.7766}$$

$$M_{big}(z_5(t)) = \frac{z_5(t) - 0.0914}{5.1006 \times 10^{-4}}, M_{little}(z_5(t)) = \frac{0.0919 - z_5(t)}{5.1006 \times 10^{-4}},$$

$$M_{big}(z_6(t)) = \frac{z_6(t) + 1.4685}{0.0433}$$

$$M_{little}(z_6(t)) = -\frac{1.4252 + z_6(t)}{0.0433}, M_{big}(z_7(t)) = \frac{z_7(t) - 0.0022}{0.0131},$$

$$M_{little}(z_7(t)) = \frac{0.0152 - z_7(t)}{0.0131}$$

$$A_1 = \begin{bmatrix} -0.0401 & 0.0401 & -0.3067 & -0.6713 \\ 0.1500 & -0.1500 & 0 & 0 \\ 87.4525 & 0 & -0.1971 & 0.0985 \\ 3.5280 & 0 & 0.0914 & -1.4252 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} -0.0401 & 0.0401 & -1.6748 & -5.6437 \\ 0.1500 & -0.1500 & 0 & 0 \\ 87.4525 & 0 & -0.2541 & 0.1271 \\ 2.7515 & 0 & 0.0919 & -1.4685 \end{bmatrix}$$

$$B_{11} = \begin{bmatrix} 0 & 0 & 0 & 0.0022 \end{bmatrix}^T \ B_{12} = \begin{bmatrix} 0 & 0 & 0 & 0.0152 \end{bmatrix}^T$$

$$C_1 = C_2 = [1 \quad 0 \quad 0 \quad 0] B_{21} = B_{22} = [0 \quad 0 \quad 0 \quad 1]^T$$

103.3.2 H∞ Controller Design

In this section, we derive a method of the H∞ controller for the NRPCS. We use the control law as follows:

Controller Rule i ***If*** $z_1(t)$ is big, $z_2(t)$ is big, $z_3(t)$ is big, $z_4(t)$ is big, $z_5(t)$ is little, $z_6(t)$ is big, and $z_7(t)$ is little, ***then***

$$u(t) = K_i x(t). \tag{103.6}$$

Then, the closed-loop system can be recast as:

$$\begin{cases} \dot{x}(t) = \sum_{i=1}^{2} \sum_{j=1}^{2} \mu_i(z(t)) \mu_j(z(t)) [(A_i + B_{1i} K_j) x(t) + B_{2i} \omega(t)] \\ y(t) = \sum_{i=1}^{2} \mu_i(z(t)) C_i x(t) \end{cases}. \tag{103.7}$$

Theorem 1 *Given* $\gamma > 0$*, if there exist the matrices* K_i *and* P *satisfying*

$$(\mathrm{E}_{ij} + \mathrm{E}_{ji}) / 2 < 0, \quad i, j = 1, 2, \tag{103.8}$$

*then System (*103.7*) is stable with an* H_∞ *performance* γ*, i.e.,* $\|y(t)\|_2 \le \gamma \|\omega(t)\|_2$ *for all* $\omega \in L_2[0, \infty)$ *under zero-initial conditions,*

where

$$\mathrm{E}_{ij} = \begin{bmatrix} A_i^T P + K_j^T B_{1i}^T P + P A_i + P B_{1i} K_j & B_{2i} & C_i^T \\ * & -\gamma^2 & 0 \\ * & * & -I \end{bmatrix}.$$

Proof We choose the Lyapunov function as

$$V(t) = x^T(t) P x(t). \tag{103.9}$$

Then, we get

$$\dot{V}(t) = \dot{x}^T(t) P x(t) + x^T(t) P \dot{x}(t). \tag{103.10}$$

Taking account of (103.7) with $\omega(t) = 0$, we can obtain

$$\begin{aligned}\dot{V}(t) &= (A_i x + B_{1i} K_j x)^T P x + x^T P (A_i + B_{1i} K_j) x. \\ &= x^T \Pi_{ij} x\end{aligned} \tag{103.11}$$

We know that (103.8) guarantees $\dot{V}(t) < 0$ for all nonzero $x(t)$. Thus, one can find a small $\upsilon > 0$ such that $\dot{V}(t) < -\upsilon \|x(t)\|^2$. The stability of the considered system is complete.

Next, assuming that $\omega(t) \neq 0$, we consider H_∞ performance of the system. We can choose the index function J as follows:

$$\begin{aligned}J &= \int_0^\infty [y^T(t) y(t) - \gamma^2 \omega^T(t) \omega(t)] dt \\ &= \int_0^\infty [y^T(t) y(t) - \gamma^2 \omega^T(t) \omega(t) + \dot{V}(x_t)] dt + V(x_t)\big|_{t=0} - V(x_t)\big|_{t=\infty}.\end{aligned} \tag{103.12}$$

Noting zero-initial conditions, $V(x_t)\big|_{t=0} = 0$, and $\omega(t) \in L_2[0, \infty)$, we obtain

$$\begin{aligned}J &\le \int_0^\infty [y^T(t) y(t) - \gamma^2 \omega^T(t) \omega(t) + \dot{V}(x_t)] dt \\ &= \int_0^\infty \sum_{i=1}^r \mu_i(z(t)) \sum_{j=1}^r \mu_j(z(t)) \begin{bmatrix} x^T(t) & \omega^T(t) \end{bmatrix} \Psi_{ij} \begin{bmatrix} x(t) \\ \omega(t) \end{bmatrix} dt\end{aligned} \tag{103.13}$$

and

$$\Psi_{ij} = \begin{bmatrix} A_i^T P + K_j^T B_{1i}^T P + P A_i + P B_{1i} K_j + C_i^T C_i & B_{2i} \\ * & -\gamma^2 \end{bmatrix} \tag{103.14}$$

Consider $\mu_i(z(t)) \ge 0, \sum_{i=1}^r \mu_i(z(t)) = 1$. Applying the Schur complement, (103.8) guarantees $J \le 0$, and we get

$$\int_0^\infty z^T(t) z(t) dt \le \int_0^\infty \gamma^2 \omega^T(t) \omega(t) dt. \tag{103.15}$$

Then, according to definition of H_∞ performance, the H_∞ performance of the system in (103.7) is established. This completes the proof.

Theorem 2 Given $\gamma > 0$, if there exist suitable matrices K_i and P satisfying

$$(\Xi_{ij} + \Xi_{ji}) / 2 < 0, \; i, j = 1, 2, \tag{103.16}$$

then System (103.7) *is asymptotically stable with an H_∞ performance γ, i.e., $\|y(t)\|_2 \le \gamma \|\omega(t)\|_2$ for all $\omega \in L_2[0, \infty)$ under zero-initial conditions. The matrices of a controller of* (103.3) *can be designed by*

$$K_j = W_j Q^{-1}, \tag{103.17}$$

where

$$\Xi_{ij} = \begin{bmatrix} QA_i^T + W_j^T B_{1i}^T + A_i Q + B_{1i} W_j & QB_{2i} & QC_i^T \\ * & -\gamma^2 & 0 \\ * & * & -I \end{bmatrix}. \tag{103.18}$$

Proof Pre and postmultiply the diagonal matrix $diag\{P^{-1}, I, I\}$ to E_{ij} in Theorem 1, and apply the manipulation of variables so that

$$Q = P^{-1}, W_j = K_j P^{-1}. \tag{103.19}$$

We can obtain Theorem 2. The proof is completed.

103.4 Simulation

Here, we use a numerical example to illustrate the advantage of the design method in this chapter.

In this case, we set the H_∞ index $\gamma = 0.1$. The H_∞ controller of the example based on the theorem we proposed is shown as follows:

$$K_1 = [-0.3375 \ -0.0002 \ 0.0019 \ -0.3804]$$
$$K_2 = [-0.6714 \ -0.0005 \ 0.0015 \ -0.7760].$$

And the other matrixes are as follows:

$$W_1 = \begin{bmatrix} -0.0072 & -0.0007 & 0.0767 & -0.0001 \end{bmatrix}$$
$$W_2 = \begin{bmatrix} -0.0136 & -0.0021 & 0.0656 & -0.0008 \end{bmatrix},$$
$$Q = \begin{bmatrix} 0.0729 & -0.0002 & 0.0455 & -0.0454 \\ -0.0002 & 16.3517 & 0.0886 & -0.0068 \\ 0.0455 & 0.0886 & 38.7363 & -0.0512 \\ -0.0454 & -0.0068 & -0.0512 & 0.0402 \end{bmatrix}$$

Generally, we choose $\omega(t) = 0.5\sin(15t)e^{-t}$, and the neutron density near to equilibrium density, $n_r = 0.15 + 0.04t$. The simulation is operating in the $t \in [0 \quad 20]$ with an initial state $[x_1, x_2, x_3, x_4] = [0, 0, 0, 0]$. Figure 103.1 shows the disturbance $\omega(t)$ and output $y(t)$.

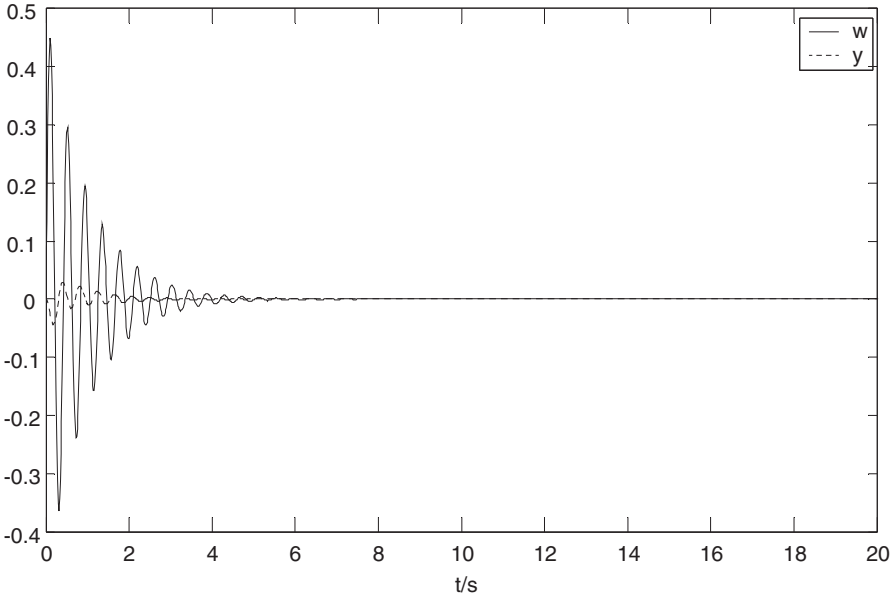


Fig. 103.1 Disturbance and the output of the closed system

From the simulation, we can know that $\|\omega(t)\|_2 = 0.25\|y(t)\|_2 = 0.022\|y(t)\|_2 / \|\omega(t)\|_2 < 0.1$, which implies that the result is effective.

103.5 Conclusion

In this chapter, the problem of H∞ of the NRPCS is considered, and an H∞ a controller design scheme is proposed. The H∞ controller can not only ensure system stability but also NRPCS performance with the energy-to-energy disturbance attenuation level. The example is used to demonstrate the advantage of the proposed method.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (61203005), the Postdoctoral Science-Research Developmental Foundation of Heilongjiang Province (LBH-Q12130), and the National Natural Science Foundation of Heilongjiang Province (QC2013C068).

References

1. Akin H, Altin V. Rule-based fuzzy logic controller for a PWR type nuclear power plant[J]. IEEE Trans Nucl Sci 1991;38(2):883–90
2. Zhou SS, Li T. Robust stabilization for delayed discrete-time fuzzy systems via basis-dependent Lyapunov–Krasovskii function [J]. Fuzzy Sets Syst. 2005;151:139–53
3. Zhang HB, Yu JB. LMI-based stability analysis of fuzzy large-scale systems with time delays [J]. Chaos, Solitons and Fractals. 2005;25:1193–207
4. Tong SC, Wang W. Decentralized control for discrete-time fuzzy large-scale systems with time-delay[C].Proceedings of the 6th World Congress on intelligent control and automation, Dalian, China, 2006; pp. 80–4
5. Edwards R, Lee K, Schultz M. State feedback assisted classical control: an incremental approach to control modernization of existing and future nuclear reactors and power plants[J]. Nucl Technol. 1990;92(2):167–83.
6. Arab-Alibeik H, Setayeshi S. An adaptive-cost-function optimal controller design for a PWR nuclear reactor[J]. Ann Nucl Energy. 2003;30(6):739–54.
7. Li CG, Wang HJ, Liao XF. Delay-dependent robust stability of uncertain fuzzy systems with time-varying delays[J].IEE Proc-Control Theory Appl. 2004;151(4):417–21.
8. Wu HN. Delay-Dependent stability analysis and stabilization for discrete-time fuzzy systems with state delay: a fuzzy Lyapunov–Krasovskii functional approach[J]. IEEE Trans Syst Man Cybern B Cybern. 2006;36(4):954–62.
9. Moon Y, Park P, Kwon W, Lee Y. Delay-dependent robust stabilization of uncertain state-delayed systems [J]. Int J Control. 2001;74(14):1447–55
10. Zhang XM, Han QL, Wu M. A new finite sum inequality for Delay-Dependent $${{H}_{\infty }}$$Control of discrete-time delay systems[C].Proceedings of the 6th World Congress on intelligent control and automation, Dalian, China; 2006.
11. Lee C. Fuzzy logic in control systems: fuzzy logic controller-part 1 and part 2[J]. IEEE Trans Syst Man Cybern. 1990;9(3):404–35.
12. Tanaka K, Wang H. Fuzzy control systems design and analysis-a linear matrix inequality approach[M]. New York: John Wiley & Sons, Inc.; 2001. pp. 39

Chapter 104
Robust Optimization Design for the Cylindrical Helical Spring Based on the Improved Particle Swarm Algorithm

Yijing Liu, Ming Liu and Minrong Zhao

Abstract High reliability is one of the most important criterions for the design of the spring; however, due to complex and diverse factors, each type of spring requires a different design formulation. In order to make full use of the capability of materials and obtain higher reliability of the cylindrical helical spring, a multiobjective optimal model is established based on robust and reliability optimization design theory. The objective was to obtain a spring with minimum mass and maximum natural vibration frequency and minimum reliability sensitivity. Some parameters of the spring are taken as random variables in this model and the reliability target is taken as the constrained condition. This multiobjective optimal problem is transformed into a single-objective optimal problem by using the multiplication and division method. Further, an improved swarm algorithm which can reasonably deal with hybrid discrete variables is used to solve this problem. The proposed method is applied to a practical example. The results show that this method is practical, reliable, and effective for obtaining parameters of robust optimization design for reliability.

Keywords Robust optimization design · Cylindrical helical spring · Particle swarm algorithm

104.1 Introduction

The cylindrical helical spring system is widely used in almost all of the mechanical and electrical products due to its structural features and elastic buffer ability. The cylindrical helical spring must be lightweight, highly reliable, and be highly resistant to fatigue to ensure that the products work well. In a traditional way, the cylindrical helical spring is designed based on empirical equations, charts, and technical

Y. Liu (✉) · M. Liu · M. Zhao
Science College, Air Force Engineering University, Changle Road, 710051 Xi'an, China
e-mail: liuyj01@sina.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_104

manuals. The weight and carrying capacity of the cylindrical helical spring depend on parameters which are chosen by subjective experience [1, 2]. The reliability of the cylindrical helical spring is estimated conservatively in this way and the capability of materials cannot bring all potentialities into full play. However, reliability is extremely design sensitive. Very slight changes to the design of parameters can cause profound changes in reliability which is why it is important to specify product reliability before any design work is undertaken. Faced with the potentially conflicting criteria of adverse performance and economy, some optimization methods are presented to obtain the optimal solution of this problem. In order to establish an optimal model, some parameters of the spring can be taken as random variables and the reliability target can be taken as the constrained condition[3, 4]. A theoretical method of reliability-based robust optimization design is proposed and a model for designing a helical compression spring is developed using multicriterial optimization to find the least change of the spring's stiffness, maximal natural frequency, and minimal spring mass [5–7]. The problem addressed here consists of optimizing mechanical properties, e.g., mass and natural vibration frequency of cylindrical helical springs given a set of design parameters, which constitutes the design space of the problem. This multiobjective optimal problem is transformed into a single-objective optimal problem by using the multiplication and division method. Further, an improved swarm algorithm is used to solve this problem which can reasonably deal with hybrid discrete variables.

104.2 Robust Optimization Reliability Design Model for Cylindrical Helical Spring

104.2.1 Mechanical Properties Measure and Design Variable

A spring is an elastic device that regains its original shape after being compressed or extended. The helical springs are made of wire with round or rectangular section wrapping up on a screw propeller. The geometric elements of a cylindrical helical spring are shown in Fig. 104.1.

The main design parameters of the cylindrical helical spring include the wire diameter of the spring d, the mean diameter D, and the number of active coils n. Its mechanical properties are listed bellow [1]:

1. Mass of the cylindrical helical spring (m). Light weight and small size have become the trend of technology development for many products. So mass is an important design parameter of the cylindrical helical spring.

$$m=\frac{\rho\pi^2 d^2 D}{4}(n+n_2) \tag{104.1}$$

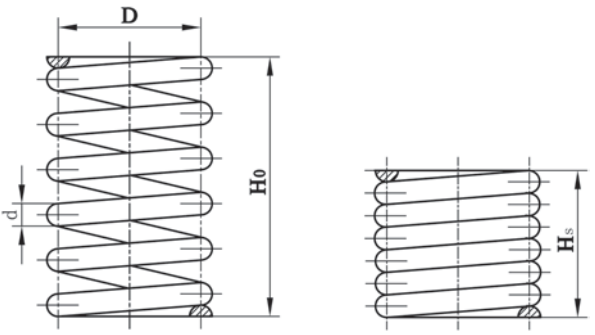


Fig. 104.1 Geometric elements of a cylindrical helical spring

where ρ is the density of the material and n_2 is the number of closed and ground coils which provide good bearing against the seats to reduce the bend effect.

2. Natural vibration frequency of the cylindrical helical spring (f). To avoid surge, the natural frequency of the cylindrical helical spring must be taken into account.

$$f=\frac{d}{2\pi nD^2}\sqrt{\frac{G}{2\rho}} \tag{104.2}$$

where G is the material shear modulus.

3. Reliability of the cylindrical helical spring (R). The failure of the spring is often caused by fracture due to fatigue. Generally fatigue strength and shear strength of the cylindrical helical spring follow normal distribution. Because of this, the reliability of the cylindrical helical spring can be calculated from the following equations

$$R=\Phi(Z_R) \tag{104.3}$$

$$Z_R=\frac{\overline{\delta_{\tau_r}}-\overline{\tau_r}}{\sqrt{\sigma_{\delta_{\tau_r}}{}^2+\sigma_{\tau_r}{}^2}} \tag{104.4}$$

where Φ(*)is the normal distribution function, $\overline{\delta}_{\tau_\Gamma}$ and $\sigma_{\delta_{\tau\Gamma}}$ are the arithmetic mean and standard deviation of limit shear fatigue strength, respectively, and $\overline{\tau}_\Gamma$ and σ_{τ_Γ} are the arithmetic mean and standard deviation of shear stress, respectively.

The wire diameter of the spring, the mean diameter, and the number of active coils are used as decision variables in this research.

$$X=[d \quad D \quad n]^T=[x_1 \quad x_2 \quad x_3]^T \tag{104.5}$$

104.2.2 Objective Function

Suppose the decision variable X is a random variable, reliability sensitivity is the partial derivative of the reliability to X. To obtain minimum mass, maximum natural frequency, and minimum reliability sensitivity, the solution of the optimization problem corresponds to the least value of the objective function. By substituting the values, objective functions can be obtained as follows:

$$\min \quad f_1(X)=\frac{\rho\pi^2 x_1{}^2 x_2}{4}(x_3+n_2) \tag{104.6}$$

$$\max \quad f_2(X)=\frac{x_1}{2\pi x_3 x_2{}^2}\sqrt{\frac{G}{2\rho}} \tag{104.7}$$

$$\min \quad f_3(X)=\sqrt{\sum_{i=1}^{3}\left(\frac{\partial R}{\partial x_i}\right)} \tag{104.8}$$

This multiobjective optimal problem is transformed into a single-objective optimal problem by using the multiplication and division method. The objective function is given by eq. 104.9.

$$\min \quad f(X)=\frac{f_1(X)f_3(X)}{f_2(X)} \tag{104.9}$$

104.2.3 Constraints

The following constrains must be satisfied to obtain optimal values of the variables.

1. Spring index. Spring index depends on the ratio D/d and is recommended to be greater or equal to 4 and less or equal to 18 [2]. If the spring index is less than 4, the spring will be difficult to manufacture and have larger stress concentration due to tight curvature.

$$g_1(X)=4-x_2/x_1\le 0 \quad \text{and} \quad g_2(X)=x_2/x_1-18\le 0 \tag{104.10}$$

2. Ratio of slenderness. In order to reduce the bend or wavy effect, the ratio of the cylindrical helical spring must be considered [2].

$$g_3(X)=1.5x_1/x_2+0.5x_3-b_c\le 0 \tag{104.11}$$

where b_c is the critical value of the slenderness ratio of the spring. When the spring is fixed on both ends, $b_c = 5.3$. When the spring is fixed on one end, $b_c = 3.7$. When the spring is not fixed on both ends, $b_c = 2.6$.

3. Natural frequency. To avoid resonance, the spring should not be cycled at a frequency close to its natural frequency.

$$g_4(X) = f_0 - \frac{x_1}{2\pi x_3 {x_2}^2}\sqrt{\frac{G}{2\rho}} \tag{104.12}$$

where f_0 is the working frequency of the spring.

4. Solid height. In order to maintain linearity when the spring is about to close, it is necessary to avoid the gradual touching of coils. Thus,

$$H_0 - \delta_{\max} \geq H_s \tag{104.13}$$

where H_0 is the free length of the spring when there is no load affecting it. $\delta_{\max}$ is the deflection when the spring is loaded the maximum external force. H_s is the height of the spring when the spring is loaded such that the coils are touching each other [1]. This constraint can be expressed as

$$g_5(X) = \frac{8F_{\max}{x_2}^3 x_3}{G{x_1}^4} + x_1 x_3 - 0.4 x_2 x_3 \leq 0 \tag{104.14}$$

5. Reliability. Suppose the allowed reliability is R_0, the reliability constraints can be obtained as

$$g_6(X) = R_0 - R \leq 0 \tag{104.15}$$

6. Spring allowable design variables. All design variables must be in the ranges allowed. Thus

$$g_7(X) = d_{\min} - x_1 \leq 0 \quad \text{and} \quad g_8(X) = x_1 - d_{\max} \leq 0 \tag{104.16}$$

$$g_9(X) = D_{\min} - x_2 \leq 0 \quad \text{and} \quad g_{10}(X) = x_2 - D_{\max} \leq 0 \tag{104.17}$$

$$g_{11}(X) = n_{\min} - x_3 \leq 0 \quad \text{and} \quad g_{12}(X) = x_3 - n_{\max} \leq 0 \tag{104.18}$$

Then the robust optimization model can be written as

$$\min \quad f(X) = \frac{f_1(X) f_3(X)}{f_2(X)} \tag{104.19}$$

$$\text{s.t.} \quad g_i(X) \le 0 \ (i = 1, 2, \ldots, 12) \tag{104.20}$$

104.3 Improved Particle Swarm Optimization (IPSO)

Particle swarm optimization (PSO) has emerged in recent years as one of the most efficient approaches to solve optimization problems. In PSO, a number of particles are placed in the search space of some problem. Each particle then determines its movement through the search space by combining some aspect of the history of its own current and best locations with those of one or more members of the swarm. The next iteration takes place after all particles have been moved. Eventually, the swarm as a whole is likely to move close to an optimum of the fitness function [8, 9]. This procedure for solving robust optimization design of the cylindrical helical spring is given in the following steps:

Step 1: Determine parameters. Suppose $X_i=[d\ D\ n]^T=[x_1\ x_2\ x_3]^T$ is the position of the particles and $V_i=(v_i1, v_i2, v_i3, v_i4)=(\Delta d, \Delta D, \Delta n)$ is the velocity of the particles. The search space is determined according to national standards in which allowable decision variables are discrete values.

Step 2: Initialization of swarm positions and velocities. Initialize a population of particles with random positions and velocities.

Step 3: Evaluation of particle's fitness. Evaluate each particle's fitness value.

Step 4: Comparison to personal best (pbest). Compare each particle's fitness with the particle's pbest. If the current value is better than pbest, then set the pbest value equal to the current value and the pbest location equal to the current location in *n*-dimensional space.

Step 5: Comparison to global best (gbest). Compare the fitness with the population's overall previous best. If the current value is better than gbest, then reset gbest to the current particle's array index and value.

Step 6: Updating of each particle's velocity and position. Change the velocity (v_i) and position of the particle (x_i) according to eq. 104.21 and 104.22

$$v_i(t+1)=\omega v_i(t) + c_1 r_1(t)(p_i(t) - x_i(t)) + c_2 r_2(t)(p_g(t) - x_i(t)) \tag{104.21}$$

$$x_i(t+1)=x_i(t) + v_i(t+1) \tag{104.22}$$

where $i = 1, 2,\ldots, N$ indicates the number of particles in the population; $t = 1, 2,\ldots, t_{max}$, indicates the iteration, ω is a parameter called the inertial weight; Positive constants c_1 and c_2 are the so-called cognitive and social components, which are the acceleration constants responsible for varying the particle velocity toward pbest and gbest. Variables r_1 and r_2 are two random values in the range [0, 1], respectively. $v_i = [v_{i1}, v_i2,\ldots, v_{in}]^T$ and $x_i = [x_i1, x_i2,\ldots, x_{in}]^T$ stand for the velocity and position of the ith particle, respectively, and $p_i = [p_i1, p_i2,\ldots, p_{in}]^T$represents the best previous position of the ith particle. P_i can be obtained by Eqs. 104.23

$$p_i(t+1) = \begin{cases} p_i(t) & if \quad f(x_i(t+1)) \geq f\left(p_i(t)\right) \\ x_i(t+1) & if \quad f\left(x_i(t+1)\right) < f\left(p_i(t)\right) \end{cases} \tag{104.23}$$

$p_g(t)$ represents the best position of the particle among all the particles in the swarm. $p_g(t) \in \{p_0(t), p_1(t),\ldots, p_s(t)\}$ and $f(p_g(t)) = \min\{f(p_0(t), p_1(t),\ldots, p_s(t))\}$.

Calculate the distance between the position of the particle (X_i) and the allowable decision variable value. Then replace X_i by allowable decision variable value with the shortest distance calculated above to make sure that the new position is in the search space.

Step 7: Checking feasible solution. Check the new position according to constraint equations. If the new position is not the feasible solution, return to Step 6 until each new position of the particle is the feasible solution of this optimization problem.

Step 8: Repeating the evolutionary cycle. Return to Step 3 until a stopping criterion is met, usually a sufficiently good fitness or a maximum number of generations.

104.4 Experimental Results

Table 104.1 shows the known conditions of an application. The valve spring is made of 50 CrVa and is fixed on both ends.

Use the optimization method proposed above, the optimization results can be obtained quickly. The mass and natural frequency are dimensionless to decrease the influence of dimensions before calculating. Table 104.2 shows the comparison of three optimization results [10].

Table 104.1 Parameters of the valve spring

F_{max} (N)	F_{max} (N)	f_0 (Hz)	G (MPa)	ρ (kg/mm^3)	d (mm)	D (mm)	n	$n2$	C	R_0
266.56	666.4	50	7.9×10^3	7.85×10^{-6}	$4\leq d\leq10$	$20\leq D\leq50$	$4\leq n\leq15$	1.75	$C\geq6$	0.999

Table 104.2 Comparison of three optimization results. (The values in the parentheses are the rounding values of calculations)

	D	*D*	*n*	*m*	*f*	*R*	*f*(*X*)
Traditional optimization design	6.0702(6)	40.05(40)	5.5	0.1881	270.56	0.999999	0.5571
Reliability optimization design	5.682(5.5)	38.199(38)	4.5	0.1390	305.34	0.999499	0.3626
Robust optimization reliability design	5.5	35	5	0.1383	323.94	0.999949	0.3400

104.5 Conclusion

The aim of this work was to optimize mass and natural vibration frequency of a cylindrical helical spring that is capable of carrying given external forces without failure and minimize reliability sensitivity. Good results were obtained with this robust optimization reliability design model proposed above. The results show that this method is practical, reliable, and effective. The mass of the cylindrical helical spring decreases and the natural vibration frequency increases significantly. In the meantime, the improved swarm algorithm solves the problem, which includes hybrid discrete variables. The future studies could focus on comparing the results with the direct use of numerical simulations in order to evaluate the accuracy of the fitness model.

References

1. Liu W. Mechanical reliability design. 2nd ed [M]. Beijing: Tshinghua Press; 1996. pp. 370–5 (In Chinese).
2. Zhang Y, Liu H, Wang D. The spring manual [M]. Beijing: Mechanical Industry Press; 1997. pp. 237–46 (In Chinese).
3. Jia Y, Zhang W. Multi-objective robust optimal design for spiral springs [J]. J Univ Sci Technol Beijing. 2009;31(6):788–91 (In Chinese).
4. Wei Z. Optimization design based on reliability for variable load cylinder helix compression spring [J]. Mach Des Manuf. 2006;25(5):28–30 (In Chinese).
5. Zhang Y, He X, Liu Q, Wen B. Reliability-based robust optimization design for automobile components part 1: theory [J]. Eng Sci. 2004;6(3):65–79 (In Chinese).
6. Liu R, Hou S, Zhang Y. Robust optimization design for reliability for coil spring based on multi-objective design marking methods [J]. J Mech Strength. 2012;34(2):214–8 (In Chinese).
7. Nugraha HS, Jauhari WA, Rosyidi CN. Optimization model for helical compression spring design of lock case with maximum reliability criteria [C]. Proceedings of Industrial Engineering and Service Science. IEEE Computer Society; 2011. pp. 255–60.
8. Coelho LdS. Particle swarm optimization combined with normative knowledge applied to Loney's solenoid design [J]. Int J Comput Math Electr Electron Eng. 2009;28(5):1155–61.
9. Poli R, Kennedy J, Blackwell T. Particle swarm optimization [J]. Swarm Intell. 2007;1(1):33–57.
10. Gong X, Liu Y, Cui L. Reliability optimization design of cylinder helical spring [J]. J Air Force Eng Univ (Nat Sci Edition). 2010;11(2):77–80 (In Chinese).

Chapter 105
Meticulous Restricted Equivalent Transformation on Singular Systems

Huanzhang Ling, Yuesheng Luo and Man Xu

Abstract In this chapter, a meticulous restricted equivalent transformation on singular systems is proposed. It can simplify different variables in complex singular systems and turn the complex singular system to be more simple and clear. This method gives the complexity of the singular system equivalent to more singular systems so that we approach the new phenomenon of altering the traditional restricted equivalent transformation provided the future analysis foundation on the practical system.

Keywords Component · Restricted equivalent transformation · Singular system · Block matrix

105.1 Introduction

There are many fields existing in the singular systems. For example, power control systems, chemical reaction processes, aerospace exploration engineering, social economic control systems, intelligent network analysis, biological control systems, time-series (discrete) analysis [1], and so on.

Generally, the continuous-time state-space description of singular systems can be expressed as below

$$\begin{cases} \mathcal{E}(v)\dot{x}(v) = \hbar\left[x(v), u(v), v\right] \\ y(v) = \tilde{\lambda}\left[x(v), u(v), v\right] \end{cases} \tag{105.1}$$

Here function $\hbar(*) \in f(x_1, x_2, \cdots, x_n)$, $\tilde{\lambda}(*) \in f(x_1, x_2, \cdots, x_m)$, and variable $v \geq 0, x(v) \in \Re^n, u(v) \in \Re^m, y(v) \in \Re^l$, $\mathcal{E}(v) \in \Re^{n\times n}, rank(\mathcal{E}(v)) = \gamma < n$.

H. Ling (✉) · Y. Luo
College of Automation, Harbin Engineering University, 150001 Harbin, China
e-mail: linghuanzhang@hrbeu.edu.cn

M. Xu
College of Science, Harbin Engineering University, 150001 Harbin, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_105

In many situations, we analyze the following continuous linear time invariant singular systems [2], which can be expressed as

$$\begin{cases} E\dot{x}(v) = Fx(v) + Tu(v) \\ y(v) = Zx(v) + Wu(v) \end{cases} \tag{105.2}$$

Here, the constant matrixes F, T, Z, W meet the conditions as below: $rank(\mathcal{E}) = \gamma < n$, $F \in \Re^{n\times n}, T \in \Re^{n\times m}, Z \in \Re^{l\times n}, W \in \Re^{l\times m}$, for the sake of convenience, the singular system (105.2) is abbreviated $(\mathcal{E}, F, T, Z, W)$.

Definition: the singular system's transfer function is $G(s)$

$$G(s) = \left\{\mathcal{E}, \begin{bmatrix} F & T \\ Z & W \end{bmatrix}\right\} \tag{105.3}$$

We can use the transformation $\overline{x} = J^{-1}x$ for singular control system (105.2), here J is a nonsingular matrix

$$\begin{cases} K\mathcal{E}J\dot{\overline{x}}(v) = KFJ\overline{x}(v) + KTJ(v) \\ y(v) = ZJ\overline{x}(v) + WJ(v) \end{cases} \tag{105.4}$$

where the matrix K and the matrix J are invertible, and the equivalent transformation of this state space singular system is constrained, which is usually limited to the equivalent transformation. Apparently, symmetry, reflexivity, and transitivity are the characteristics of the singular system-constrained equivalent transformation.

105.2 Description of the Problem

Given the regular matrix $(\mathcal{E}, F)$, $\mathcal{E}, F \in \Re^{n\times n}$, there are two invertible matrixes K, J,

$$K\mathcal{E}J = \begin{bmatrix} I_\gamma & 0 \\ 0 & \mathcal{H} \end{bmatrix}, KFJ = \begin{bmatrix} K_1 & 0 \\ 0 & I_{n-\gamma} \end{bmatrix},$$

$$KT = [T_1 \mid T_2], ZJ = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix} \tag{105.5}$$

where the nilpotent matrix $\mathcal{H} \in \Re^{(n-\gamma)\times(n-\gamma)}, \hbar = \min\{\mid \mathcal{H}^v = 0, v \in \mathcal{Z}^+\}$, let $J^{-1}x(v) = [x_1(v) \mid x_2(v)], x_1(v) \in \Re^r, x_2(v) \in \Re^{n-r}$, we get

$$\begin{cases} \dot{x}_1(\nu) = K_1x_1(\nu) + T_1u(\nu) \\ \mathcal{H}\dot{x}_2(\nu) = x_2(\nu) + T_2u(\nu) \\ y_1(\nu) = Z_1x_1(\nu), y_2(\nu) = Z_2x_1(\nu) \\ y(\nu) = y_1(\nu) + y_2(\nu) \end{cases} \tag{105.6}$$

This is the first restriction equivalent form, which is referred to the fast and slow decomposition of the subsystem. If $\dot{x}_1(v)$ is the slow subsystem [3, 4], $\mathcal{H}\dot{x}_2(v)$ is the quick subsystem [5], the transfer function is

$$G(s) = Z_1(sI_\gamma - K_1)^{-1}T_1 + Z_2(s\mathcal{H} - I_{n-\gamma})^{-1}T_2 \tag{105.7}$$

105.3 Results

Theorem linear time-invariant singular system mode (2), suppose $W=0$, matrixes $K, J \in \Re^{n\times n}$, $rank(K)=rank(J)=n$, let $x=\mathcal{J}\bar{x}$, we get the singular system transformation expressed as below

$$\begin{bmatrix} I_{11} & & & & & & & \\ & I_{22} & & & & & & \\ & & I_{r1} & & & & & \\ & & & I_{r2-r1} & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1(v) \\ \dot{x}_2(v) \\ \dot{x}_3(v) \\ \dot{x}_4(v) \\ \dot{x}_5(v) \\ \dot{x}_6(v) \\ \dot{x}_7(v) \\ \dot{x}_8(v) \end{bmatrix} = \begin{bmatrix} F_{11} & O & F_{13} & O & O & O & O & O \\ F_{21} & O & F_{23} & O & O & O & O & O \\ O & O & O & O & I_{r-r_1} & O & O & O \\ O & O & O & O & O & I_{r_1} & O & O \\ O & I_{u_1} & O & O & O & O & O & O \\ O & O & O & I_{r_1} & O & O & O & O \\ O & O & O & O & O & O & O & O \\ O & O & O & O & O & O & O & I_{s_1} \end{bmatrix}$$

$$\begin{bmatrix} x_1(v) \\ x_2(v) \\ x_3(v) \\ x_4(v) \\ x_5(v) \\ x_6(v) \\ x_7(v) \\ x_8(v) \end{bmatrix} + \begin{bmatrix} 0 & T_{11} & T_{12} & T_{13} \\ 0 & T_{21} & T_{22} & T_{23} \\ 0 & T_{31} & T_{32} & T_{33} \\ 0 & T_{41} & T_{42} & T_{43} \\ 0 & \begin{bmatrix} 0 & 0 \\ c_1 & d_1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ c_2 & d_2 \end{bmatrix} & \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ 0 & \begin{bmatrix} 0 & 0 \\ c_3 & d_3 \end{bmatrix} & \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} & 0 \\ \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} & 0 & 0 & 0 \\ 0 & \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1(v) \\ u_2(v) \\ u_3(v) \\ u_4(v) \end{bmatrix} \tag{105.8}$$

where $\dot{x}(v) = [\dot{x}_1(v) \mid \dot{x}_2(v) \mid \dot{x}_3(v) \mid \dot{x}_4(v) \mid \dot{x}_5(v) \mid \dot{x}_6(v) \mid \dot{x}_7(v) \mid \dot{x}_8(v)]$

So we can get subsystems as below

$$\dot{x}_1(v) = F_{11}x_1(v) + F_{13}x_3(v) + T_{11}u_2(v) + T_{12}u_3(v) + T_{13}u_4(v) \tag{105.8.1}$$

$$\dot{x}_2(v) = F_{21}x_1(v) + F_{23}x_3(v) + T_{21}u_2(v) + T_{22}u_3(v) + T_{23}u_4(v) \tag{105.8.2}$$

$$\dot{x}_3(v) = x_5(v) + T_{31}u_2(v) + T_{32}u_3(v) + T_{33}u_4(v) \tag{105.8.3}$$

$$\dot{x}_4(v) = x_6(v) + T_{41}u_2(v) + T_{42}u_3(v) + T_{43}u_4(v) \tag{105.8.4}$$

$$0 = x_2(v) + \begin{bmatrix} 0 & 0 \\ c_1 & d_1 \end{bmatrix} u_2(v) + \begin{bmatrix} 0 & 0 \\ c_2 & d_2 \end{bmatrix} u_3(v) + \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} u_4(v) \tag{105.8.5}$$

$$0 = x_4(v) + \begin{bmatrix} 0 & 0 \\ c_3 & d_3 \end{bmatrix} u_2(v) + \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} u_3(v) \tag{105.8.6}$$

$$0 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} u_1(v) \tag{105.8.7}$$

$$0 = x_8(v) + \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} u_2(v) \tag{105.8.8}$$

Proof there are invertible matrixes K_1 and J_1, $K_1\mathcal{E}J_1 = diag(I_q, 0)$,

Let $J_1^{-1}x(v) = [\tilde{x}_1(v) \mid \tilde{x}_2(v)]$, $\tilde{x}_1(v) \in \Re^q, \tilde{x}_2(v) \in \Re^{n-q}$, the restricted equivalent form is

$$\begin{cases} \dot{\tilde{x}}_1(v) = F_{11}\tilde{x}_1(v) + F_{12}\tilde{x}_2(v) + \tilde{T}_1 u(v) \\ 0 = F_{21}\tilde{x}_1(v) + F_{22}\tilde{x}_2(v) + \tilde{T}_2 u(v) \\ y(v) = \tilde{Z}_1\tilde{x}_1(v) + \tilde{Z}_2\tilde{x}_2(v) \end{cases} \tag{105.8.9}$$

where $K_1\mathcal{E}J_1 = \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix}$, $K_1J_1 = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$, $K_1T = [\tilde{T}_1 \mid \tilde{T}_2]$, $ZJ_1 = [\tilde{Z}_1 \;\; \tilde{Z}_2]$.

Then gradually with simultaneous diagonalization of matrices $[\mathcal{E}, F]$, we have invertible matrixes $[K_2, J_2], [K_3, J_3], \supset, [K_7, J_7]$,we get formula (105.8).

105.3.1 Comments

1. Since there was no $x_7(v)$, we can assume that there is no constraint on the system, then we can call it independence of the state of the system;
2. From 105.8.8,

$$x_8(v) = -\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} u_2(v) = -\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{u}_2(v) \\ \hat{u}_2(v) \end{bmatrix} = -\hat{u}_2(v) \tag{105.8.10}$$

$x_8(v)$ will be completely determined by the upper half of the $u_2(v)$. When other equations do not contain $x_8(v)$, there is no affliction of the state of the systems;

3. From 105.8.6,

$$x_4(v) = -\begin{bmatrix} 0 & 0 \\ c_3 & d_3 \end{bmatrix} u_2(v) - \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} u_3(v) = -\begin{bmatrix} 0 & 0 \\ c_3 & d_3 \end{bmatrix} u_2(v) - \hat{u}_3(v) \tag{105.8.11}$$

$x_4(v)$ will be completely determined by the input $u_2(v)$ and upper half of the $u_3(v)$, however other equations contain $x_4(v)$; so it will affect other states;

4. From 105.8.5,

$$\begin{aligned} x_2(\nu) &= -\begin{bmatrix} 0 & 0 \\ c_1 & d_1 \end{bmatrix} u_2(\nu) - \begin{bmatrix} 0 & 0 \\ c_2 & d_2 \end{bmatrix} u_3(\nu) - \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} u_4(\nu) \\ &= -\begin{bmatrix} 0 & 0 \\ c_1 & d_1 \end{bmatrix} u_2(\nu) - \begin{bmatrix} 0 & 0 \\ c_2 & d_2 \end{bmatrix} u_3(\nu) - \hat{u}_4(\nu) \end{aligned} \tag{105.8.12}$$

$x_2(v)$ will be completely determined by the input $u_2(v)$, $u_3(v)$and the upper half of the $u_4(v)$, however other equations contain $x_2(v)$, so it will affect other states also;

5. From 105.8.4

$$x_6(v) = -\begin{bmatrix} 0 & 0 \\ c_3 & d_3 \end{bmatrix} \dot{u}_2(v) - \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \dot{u}_3(v) - T_{41}u_2(v) - T_{42}u_3(v) - T_{43}u_4(v) \tag{105.8.13}$$

$x_6(v)$ will be completely determined by the input $u_2(v), u_3(v), u_4(v)$, and the upper half of the $u_2(v)$ and $u_3(v)$;

6. The original system can be determined by the following singular system of relationships between the state of the above system components,

$$\begin{cases} \dot{x}_1(v) = F_{11}x_1(v) + F_{13}x_3(v) + T_{11}u_2(v) + T_{12}u_3(v) + T_{13}u_4(v) \\ 0 = F_{21}x_1(v) + F_{23}x_3(v) + T_{21}u_2(v) + T_{22}u_3(v) + T_{23}u_4(v) \\ \quad + \begin{bmatrix} 0 & 0 \\ c_1 & d_1 \end{bmatrix} \dot{u}_2(v) + \begin{bmatrix} 0 & 0 \\ c_2 & d_2 \end{bmatrix} \dot{u}_3(v) + \dot{\hat{u}}_4(v) \\ \dot{x}_3(v) = x_5(v) + T_{31}u_2(v) + T_{32}u_3(v) + T_{33}u_4(v) \end{cases} \tag{105.8.14}$$

That is

$$\begin{aligned} \begin{bmatrix} I & & \\ & I & \\ & & 0 \end{bmatrix} \begin{pmatrix} \dot{x}_1(v) \\ \dot{x}_3(v) \\ \dot{x}_5(v) \end{pmatrix} &= \begin{bmatrix} F_{11} & F_{13} & 0 \\ 0 & 0 & I \\ F_{21} & F_{23} & 0 \end{bmatrix} \begin{pmatrix} x_1(v) \\ x_3(v) \\ x_5(v) \end{pmatrix} + \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{31} & T_{32} & T_{33} \\ T_{21} & T_{22} & T_{23} \end{bmatrix} \begin{bmatrix} u_2(v) \\ u_3(v) \\ u_4(v) \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 \\ \begin{bmatrix} 0 & 0 \\ c_1 & d_1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ c_2 & d_2 \end{bmatrix} & \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{u}_2(v) \\ \dot{u}_3(v) \\ \dot{u}_4(v) \end{bmatrix} \end{aligned} \tag{105.8.15}$$

Obviously (105.8.15) remains a singular system, which we still can be used in front of the decomposition theorem for system structural analysis, and then get more

detailed components of the relationship between the state and the state of the input; so the study complex high-dimensional singular system will be of great help.

105.4 Conclusion

This study acquires a meticulous restricted equivalent transformation on the state-space equation of the singular linear control system from this theorem, and reached the standard form of system architecture by the theory of matrix decomposition and transformation. This theorem can simplify a complex singular system into a standard and meticulous form; according to this form, the classifications of the states are proposed and the complex singular system problems have become much simpler.

Acknowledgments Heartfelt appreciation to the thoughtful suggestions to this chapter; and it is sponsored by the National Natural Science Foundation of China (Grant No: 51309068).

References

1. Chen C. The Meaning and Properties of Singular System. J Guangdong Polytechn Norm Univ. 2001;(4):1–8.
2. Luenberger DG, Arbel A. Singular dynamical Leontief systems. Econometrica. 1977;45(4): 991–5.
3. Lu G, Ho DWC. Continuous stabilization controllers for singular bilinear systems: the state feedback case. Automatica. 2006;42(2):309–14.
4. Duan G-R. Analysis and design of descriptor linear systems. Vol. 23. New York: Springer; 2010. pp. 35–41.
5. Misra P, Van Dooren P, Varga A. Computation of structural invariants of generalized state-space systems. Automatica. 1994;30(12):1921–36.

Chapter 106
Intelligent Control of a Servo-Motor-Driven Shock Absorber Performance Tester

Yujian Wang, Yongchang Du and Pu Gao

Abstract This chapter describes a project on the driving and controlling system of an automotive shock absorber tester. The system structures and key parameters are designed followed by the analysis of function partitioning and collaboration of a host personal computer (PC) and the controller. The communication protocol between controller and host PC is developed. Then, the structure and the flow chart of the controller program are presented. Finally, the performance test result using the tester with the developed drive system is given and compared with a qualified tester, which shows the accuracy of the new system.

Keywords Shock absorber · Performance tester · Servo controller · Programming

106.1 Introduction

The performance tester is an important facility in shock absorber and relevant industries. Its test result is a major guideline for research and development of a new product and provides an essential quality index for manufacturers [1, 2].

The tester with a frequency variable-driven eccentric is widely used in manufactures [3]. Although it is simple and cheap, its drawbacks, such as inaccurate exciting curve and incapability of a variable stroke test, limit its use. The electro-hydraulic servo tester can solve the abovementioned problems, but it has a very high purchasing/maintenance cost and large space occupation. As a result, it is more used in laboratories and high-end shock absorber manufactures [4–7].

The electronic servo techniques have been widely used in industry since the 1990s. It combines the advantage of servo control with compact size, low cost

Y. Du (✉) · Y. Wang
State Key Laboratory of Automotive Safety and Energy, Tsinghua University,
100084 Beijing, China
e-mail: duyc@tsinghua.edu.cn

P. Gao
School of Vehicle and Energy Engineering, Yanshan University, 066004 Hebei, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_106

and high performance. This chapter presents parts of efforts in the research and development of a shock absorber performance tester driven by an electronic servo actuator. Its power source is a permanent magnet (PM) synchronous servo motor, which drives a precision ball screw assembly to convert the angular movement of the motor rotor to the linear vibration of an exciting piston. The motor is controlled by an intelligent servo controller, which is responsible for the motion control, process control and communication with the host PC. The principle diagram and key parameters of the system are listed. The function partitioning and collaboration of host PC and controller is analysed and the communication protocol is developed. Then, the flow chart of the controller program is presented. Finally, the performance test result using the tester with the developed drive system is given and compared with that of traditional testers with the results indicating the advancement of the new system.

106.2 System Diagram and Key Parameters

The drive system is composed of a servo controller, servo motor, synchronous belt and pulley and ball screw (Fig. 106.1). As shown in Fig. 106.1, the servo controller plays the key role in the system. It communicates with the host PC through a controller area network (CAN) interface, generates motor motion signals according to test specifications and controls the motor to realize the motion. But it should be noted that sensors for the measurement of force, displacement, temperature and rotating speed and actuators such as clutch and fixture are not shown in the picture.

The key parameters of the system are listed in Table 106.1.

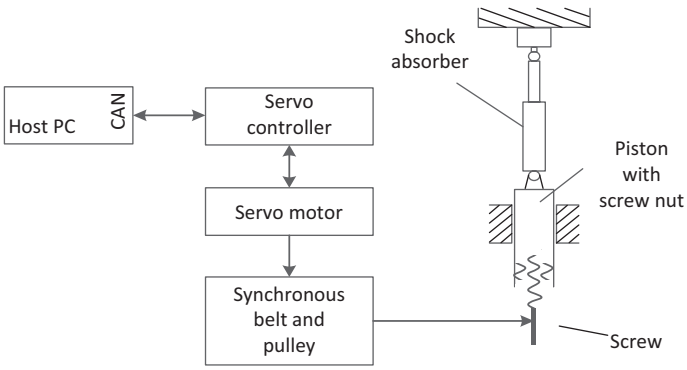


Fig. 106.1 System diagram

Table 106.1 Key parameters

Component	Feature name	Feature
Servo controller	Motor control	Speed/torque actuating drive Electronic cam
	Capacity	Rated output: 59A
	Programming	Graphic programming
	Communication	Integrated CAN open
Servo motor	Type	Brushless, PM synchronous motor
	Capacity	Rated power/speed: 29 KW/209 rad/s
	Encoder	Incremental rotary
Screw assembly	Screw type	Precise ball screw
	Lead	50 mm
	Max. stroke	200 mm

Table 106.2 Test parameters

Name	Address	Type	Format
Number of curve	C3001/0	Unsigned long	1–16
Test speed	C3002/0~C3009/0 C3029/0~C3036/0	Unsinged int	16,384/1.5 m/s
Test stroke	C3012/0~C3019/0 C3021/0~C3028/0	Unsinged int	10,000/+50 mm
Home position	C2643	Signed long	−10,000/0.1 mm
Home speed	C2644	Signed long	10,000/0.1 mm/s
Moving target	C3037	Signed long	10,000/0.1 mm
Moving speed	C5053	Signed long	10,000/0.1 mm/s

106.3 Communication Protocol

The host PC's tasks include providing graphical user interface (GUI), manipulating test data and generating test report. The intelligent servo controller is a robust hardware running an embedded operating system and is capable of real-time tasks. It has enough I/O ports and is user-programmable. Thus, in this system, the controller is responsible for generating the motion curves and controlling the motor to realize the curve.

The communication between the controller and host PC follows the CAN open protocol. It is a typical master–slave model. The host PC is responsible for initiating a request. The controller is always waiting for the request and sends back the requested data to the host PC immediately. A set of communication protocol is developed, including: (1) download the test parameters, (2) parse the running commands and (3) retrieve the current status. Details of the communication protocols are shown in Tables 106.2–106.4.

Table 106.3 Control command

Name	Address	Type	Format
Control command	C3040/0	Unsigned int	0×0001: go home 0×0002: start testing 0×0004: stop 0×0008: move 0×0010: reset

Table 106.4 Status parameters

Name	Address	Type	Format
Current status	C3041/0	Unsigned int	0×0001: going home 0×0002: testing 0×0004: moving 0×0008: error 0×0010: hardware disabled 0×0020: exceeding upper limit 0×0040: exceeding lower limit
Current position	C2575/0	Signed long	10,000/0.1 mm
Current curve no.	C3042/0	Signed long	Only valid when status is "testing"
Hardware ready	C444/1	Signed int	1 = hardware ready
CAN ready	C444/2	Signed int	1 = CAN module ready

106.4 Programming of the Controller

The controller is programmed to realize the required complex functions and process. The self-defined position follower function block is responsible for generating the exciting curves. The shock absorber test standard requires the exciting position be sinusoidal [7], but such exciting curves may result in mechanical shocks when the adjacent test velocities are different.

In this research, the theoretic sinusoidal signals are processed at the beginning and ending part of the signal. As an example, the beginning 30° rotation is stretched to 60°, and the position expression is

$$x = -A\sin[\omega(c_1 t^7 + c_2 t^6 + c_3 t^5 + c_4 t^4)] \tag{106.1}$$

where A is the test stroke and $c_1 - c_4$ are time independent coefficients.

As seen, the x and its first-, second-, third- and fourth- order derivatives are all zero at the beginning ($t = 0$). So whether the adjacent test speeds are different or equal, the position, velocity and acceleration will change smoothly. No shock will appear. The result curves are shown in Fig. 106.2.

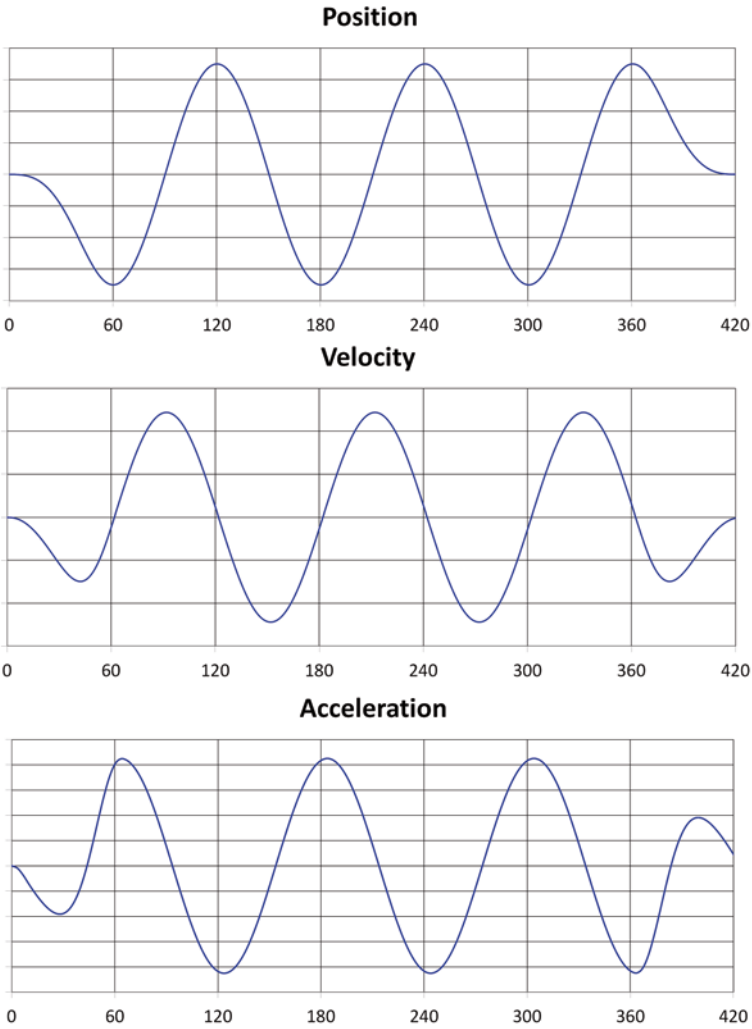


Fig. 106.2 Exciting position, velocity and acceleration curve

106.5 Validations

In addition to the programing of the servo controller, the test client software running on host PC is also developed. Figure 106.3 shows a typical result curve of a double-tube front shock absorber for a passenger car. The test sample is kept in good condition and its damping forces have been tested on a qualified electro-hydraulic servo tester. It can be seen that the curve is smooth and well filled. This means the

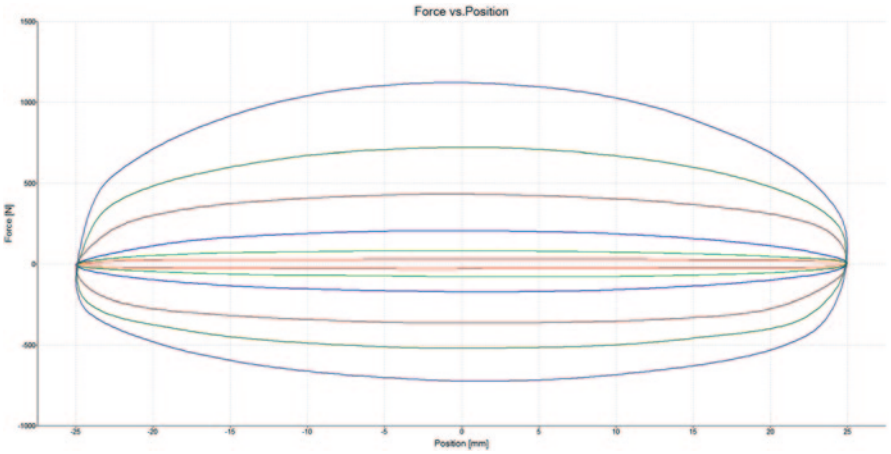


Fig. 106.3 A typical result curve

tester runs smoothly. The developed control system can provide the required exciting movements and avoid mechanical shocks.

Then, the damping force results from two machines are compared. The maximum relative error between corresponding test speeds is less than 2 %. Considering the repetitive variance of shock absorbers, this error is acceptable. The accuracy of the developed tester is good enough for the shock absorber performance tests (Table 106.5).

Table 106.5 Test results

SEq.	Speed (mm/s)	Stretching force (N)	Compressing force (N)
1	19.9	27.6	−35.4
2	49.9	78.4	100.0
3	99.9	191.1	203.0
4	299.8	401.9	362.1
5	600.9	677.0	525.4
6	999.3	1065.0	738.0

106.6 Conclusion

1. The servo-motor-driven shock absorber performance tester combines the advantages of a servo control with compact size, low cost and high performance. The intelligent control system is the core of the servo-motor-driven shock absorber performance tester. Details of the development process are provided.
2. The sample test results show that the developed tester fulfils the requirements of the shock absorber performance test specification and achieves high accuracy.

References

1. Benaziz M, Nacivet S, Deak J, et al. Double tube shock absorber model for noise and vibration analysis. SAE technical paper 2013-01-1912, 2013.
2. Czop P, SŁawik D. A high-frequency first-principle model of a shock absorber and servo-hydraulic tester. Mech Syst Signal Process. 2011;25(6):1937–55.
3. Dixon J. The shock absorber handbook. 2nd Edn. USA: Wiley; 2007 pp. 122–218.
4. Gajda J, Sidor T. Optimal choice of samples number for fast numerical velocity determination in microcomputer driven shock-absorber tester. Syst Anal Model Simul. 1995;17(2):129–33.
5. Kruse A, Gross A, Biermann J W. Analysis and optimization of shock absorber behavior in electric vehicles. SAE technical paper 2013-01-1236, 2013.
6. Nozaki H, Inagaki Y. Technology for measuring and diagnosing the damping force of shock absorbers and the constant of coil springs when mounted on a vehicle. JSAE Rev. 1999;20(3):413–9.
7. QC/T 545–1999. Technology conditions of automotive shock absorber bench test methods. Chinese national standard for automotive industry. (In Chinese).

Chapter 107
Model Predictive Controller and Fractional-Order Proportional-Integral Controller Application in the Cascade System

Dong Wu and Lei Dong

Abstract The model predictive controller (MPC) can deal with the pure delay system directly and has a good tracking performance. It also has strong robustness when model mismatch happens. The fractional-order proportional-integral (FOPI) controller has a fast response and disturbance rejection. A new MPC-FOPI controller is proposed to deal with the delay and interference links in the cascade system. The new controller is designed to control the cascade system. The results show that the new controller has the advantages of the two mentioned controllers and can effectively achieve the cascade system control.

Keywords MPC · FOPI · Delay system · Cascade system

107.1 Introduction

The model predictive controller (MPC), as a predictive control algorithm which has been widely applied and valued in industrial process control in recent years [1], can deal with the pure delay system directly and has a good tracking performance. It also has a strong robustness when model mismatch happens; however, as the dynamic matrix control (DMC) has a large sample period, the antiinterference ability of the MPC is not as good as the traditional proportional-integral-derivative (PID) controller. Furthermore, the feedback link cannot take account of both the robustness and antiinterference due to the single-level control structure of the MPC. Thus, how to deal with the relationship of the robustness and antiinterference is still a big problem of designing a controller.

D. Wu (✉) · L. Dong
School of Automation, Beijing Institute of Technology, 100081 Beijing, China
e-mail: wjireturn@163.com

L. Dong
e-mail: Pemc.bit@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_107

The dual-vessel water tank and the unmanned aerial vehicle (UAV) attitude are common in a cascade system, in which the UAV attitude control requires accuracy and it shows that the cascade PID control can meet the requirement [2]. As a common industrial process object, the boiler drum level [3] is a cascade system which has large inertial and some time-varying parameters. The control strategy of MPC-MPC can meet the requirement [4]. A continuous pharmaceutical tablet manufacturing process via direct compaction is a system which has the feature of large time-delayed, time-varying and nonlinear in operating conditions. It shows that the MPC-PID control can have a better performance than the cascade PID control [5, 6]. In other cascade systems in the industrial process control, the DMC-PID is also widely used [7]. All these applications have proved that the MPC-PID controller can deal with a large-delay system effectively [8].

As the traditional PID controller is inflexible and has oscillation and large overshoot sometimes, a new $PI^{-\lambda}D^{\mu}$ controller occurs, which has two more parameters than the PID controller; so it expands the tuning range of the controller parameters. Under this circumstance, it is a tendency that a new controller, which combines the advantages of the MPC controller and the $PI^{-\lambda}D^{\mu}$, controller is to be designed [9]. The rapid system is regulated by the $PI^{-\lambda}D^{\mu}$ controller, and the big delay system is regulated by the MPC controller. The simulations show that it has a better control effect than the PI-PI controller and DMC-PID controller.

107.2 The Concept of the MPC and FOPI Controller

107.2.1 Establishment of a No-Delay Prediction Model of State Space

Assuming that the state vector is measured, that is $\hat{\mathrm{x}}(k \mid k) = x(k)$, the effect of disturbance and measurement noise on the system can be ignored at this time. Now, the iteration model can be used to realize the future process prediction.

$$\begin{cases} x(k+1) = Ax(k)+Bu(k) \\ y(k) = C_y x(k) \\ Z(k) = C_z x(k) \end{cases}, \tag{107.1}$$

where x is the state vector; u is the control input vector; y represents the measured output; z is the controlled output vector. y and z may be overlapped or the same (when all of the controlled output is measurable. A, B and C are constant.

Write it in the matrix form:

$$x(k+j|k) = \mathrm{A}^{j} x(k)+\begin{bmatrix} A^{j-1} & A^{j-2} & \cdots & I \end{bmatrix} B \begin{bmatrix} u(k|k) \\ \vdots \\ u(k+j-1|k) \end{bmatrix}, j = 1,\cdots N. \tag{107.2}$$

The first line use $u(k\,|\,k)=u(k)$ instead of $u(k)$ because the value of $u(k)$ cannot be determined or detected at the moment k. Considering $\Delta u(k+j\,|\,k)=u(k+j\,|\,k)\text{-}u(k+j01\,|\,k)$, the value of $u(k\text{-}1)$ is known at the moment k and the matrix can be written as follows:

$$\begin{cases} u(k\,|\,k)=\Delta u(k\,|\,k)+u(k\text{-}1) \\ u(k+1\,|\,k)=\Delta u(k+1\,|\,k)+\Delta u(k\,|\,k)+u(k\text{-}1) \\ \vdots \\ u(k+N_C\text{-}1\,|\,k)=\Delta u(k+N_C\text{-}1\,|\,k)+\cdots+\Delta u(k\,|\,k)+u(k\text{-}1). \end{cases} \tag{107.3}$$

At the moment k, the predict input at the moment $k+N_c$ is

$$\begin{aligned} x(k+N_C\,|\,k)=A^{N_C-1}x(k)+(A^{N_C-1}+\cdots+A+I)B\Delta u(k\,|\,k)+\cdots \\ +B\Delta u(k+N_C-1\,|\,k)+(A^{N_C-1}+\cdots+A+I)Bu(k\text{-}1)\,. \end{aligned} \tag{107.4}$$

Finally, take the results above into formula (107.1). write it in matrix form:

$$x(k+j\,|\,k)=A^j x(k)+\left[\sum_{i=0}^{j01}A^iB\quad\cdots\quad\sum_{i=0}^{j0N_C}A^iB\right]\begin{bmatrix}\Delta u(k\,|\,k)\\ \vdots \\ \Delta u(k+N_C\text{-}1\,|\,k)\end{bmatrix}+\sum_{i=0}^{j01}A^iBu(k\text{-}1). \tag{107.5}$$

The model prediction equation of y is:

$$y(k+j|k)=Cx(k+j|k)\,j=1,\cdots,N. \tag{107.6}$$

107.2.2 Concept of Fractional Operator and the Fractional Order Controller

Define the fractional calculus operator ${}_{\alpha}D_t^{\alpha}$ [10]:

$${}_aD_t^{\alpha}(f(t))=\frac{1}{\Gamma(n\text{-}\alpha)}\frac{\mathrm{d}^n}{\mathrm{d}t^n}\int_a^t\frac{f(t)}{(t\text{-}\tau)^{\alpha\text{-}n\text{-}t}}\mathrm{d}\tau, \tag{107.7}$$

$n\text{-}1<\alpha<n+1$, $\Gamma()$ is the Gamma function

$$\Gamma(\alpha)=\int_0^{\infty}v^{\alpha-1}e^{-v}dv. \tag{107.8}$$

When $\alpha<0$, ${}_{\alpha}D_t^{\alpha}$ means the fractional integral calculus operator. When $\alpha>0$, it means the fractional differential calculus operator. Another two parameters of differential order μ and integral order λ are added to be compared with the traditional PID controller. The overshoot of system is mainly effected by μ and λ; and the steady-state

precision of the system is mainly effected by λ [11]. With the two parameters μ and λ, the FOPI controller can be designed more flexible than the traditional controller.

The fractional order controller is:

$$u(t) = k_p e(t) + k_i D_t^{-\lambda} e(t) + k_d D_t^{\mu} e(t). \tag{107.9}$$

The fractional order $\lambda \in [0,1], \mu \in [0,1]$, K_p, K_i, K_d are the parameters of the $PI^{-\lambda}D^{\mu}$ controller. The transfer function is:

$$G(s) = k_p + k_i s^{-\lambda} + k_d s^{\mu}. \tag{107.10}$$

107.2.3 Cascade-Coupling System

The series-coupling system is often found in the industry. It generally contains a fast-response link and a large-delay link so that every input and output are mutually connected. This kind of system is very difficult to control by the traditional PID controller. As shown in Fig. 107.1, U1 and U2 are the input of this system, and Y1 and Y2 are the output of this system. Y1 contacts U1 by a transfer function G1, and it contacts U2 by a transfer function G3, and Y2 contacts U1 by a transfer function G2, and it contacts U2 by a transfer function G4 when this system is decoupled. As shown in Fig. 107.2, the decoupled transfer function G1 can be treated as a cascade system which is combined by a fast-response link and a large-delay link.

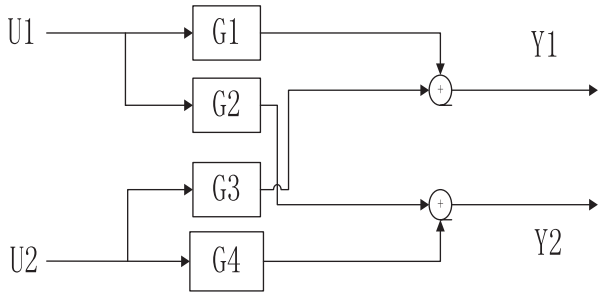


Fig. 107.1 Cascade-coupling system

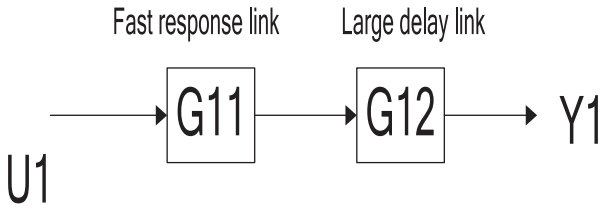


Fig. 107.2 The structure of G1

107.3 MPC-FOPI Controller Design

107.3.1 Process of the MPC Controller Design

Firstly, predict the current real output value in the measurable disturbances:

$$\hat{y}_m(k) = C_m \hat{x}(k \mid k-1) + D_{vm} v(k). \tag{107.11}$$

Secondly, use the Kalman filter array M to estimate the true state:

$$\begin{bmatrix} \hat{x}(k \mid k) \\ \hat{x}_d(k \mid k) \\ \hat{x}_m(k \mid k) \end{bmatrix} = \begin{bmatrix} \hat{x}(k \mid k-1) \\ \hat{x}_d(k \mid k-1) \\ \hat{x}_m(k \mid k-1) \end{bmatrix} + M(y_m(k) - \hat{y}_m(k)). \tag{107.12}$$

Thirdly, use the optimal state estimation for manipulated variables (MV) increment quadratic programming (QP):

$$J_k[x(k),u] = \sum_{j=1}^{N} \left\| y(\mathrm{k+j} \mid \mathrm{k}) - y_{\mathrm{ref}}(k+j \mid k) \right\|^2_{Q(\mathrm{j})} + \sum_{j=0}^{N_C-1} \left\| \Delta u(k+j \mid k) \right\|^2_{R(\mathrm{j})}. \tag{107.13}$$

Finally, predict the next moment state $\hat{x}(k+1 \mid k)$:

$$\begin{bmatrix} \hat{x}(k+1 \mid k) \\ \hat{x}_d(k+1 \mid k) \\ \hat{x}_m(k+1 \mid k) \end{bmatrix} = \begin{bmatrix} A\hat{x}(k \mid k) + B_u u(k) + B_v v(k) + B_d \quad \bar{C}\hat{x}_d(k \mid k) \\ \bar{A}\hat{x}_d(k \mid k) \\ \tilde{A}\hat{x}_m(k \mid k) \end{bmatrix}. \tag{107.14}$$

107.3.2 Design of the MPC-FOPI Controller

As shown in Fig. 107.3, G1 is composed by the decoupled transfer function and an FOPI controller. G2–G4 is in the same situation. The MPC controller is the outer loop controller. The inputs are U1 and U2, and the outputs are Y1 and Y2.

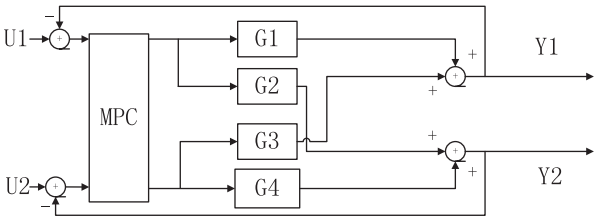


Fig. 107.3 The structure ot the MPC-FOPI controller

Firstly, use the FOPI controller to stabilize the fast response link (or inner link) in G1–G4. The parameters of the FOPI are tuned by an intelligent algorithm, such as PSO or BFOA algorithm.

Secondly, based on the cascade structure formed by the closed-loop transfer function of the inner link, the large-delay link (or outer link) builds up the MPC controller.

107.4 Simulation

Control a single-cascade system with no interference. Build up the four controllers PI-PI, FOPI-FOPI, MPC-PI and MPC-FOPI, respectively, and control the single-cascade system $G_{\mathrm{IN}}=\dfrac{-5.72}{56s+1}$ $G_{\mathrm{OUT}}=\dfrac{8e^{-2s}}{89s+1}$. The reference trajectory is $r=1$. Use integral of time-weighted absolute error (ITAE) performance index $J=\int_0^t |e(t)|tdt$ as a basis for evaluation. Use intelligent algorithm gets the optimal controller.

The inner loop PI controller and the outer loop PI controller are, respectively,

$$\mathrm{PI}_{\mathrm{IN}}=0.4086-\frac{0.0005}{\mathrm{s}},\mathrm{PI}_{\mathrm{OUT}}=0.0873+\frac{0.0007}{\mathrm{s}}. \tag{107.15}$$

The inner FOPI controller and the outer loop FOPI controller are, respectively,

$$\mathrm{PI}_{\mathrm{IN}}^{-\lambda}=0.0327+0.0234\mathrm{s}^{-0.2607},\mathrm{PI}_{\mathrm{OUT}}^{-\lambda}=0.0113+0.0272\mathrm{s}^{-0.1912}. \tag{107.16}$$

The performance index, respectively, is

$$\begin{aligned} J_{\text{PI-PI}}&=97.1796 \quad J_{\text{FOPI-FOPI}}=51.7475\\ J_{\text{MPC-PI}}&=35.1646 \quad J_{\text{MPC-FOPI}}=32.7371. \end{aligned} \tag{107.17}$$

In Fig. 107.4, the MPC-FOPI has the best performance when a cascade short-time delay system is controlled, and the MPC-PI controller can also have a better performance than the traditional PI-PI controller. It is shown that the MPC controller has a good performance when a system with delay is controlled, and the FOPI-FOPI controller has faster response ability than the PI-PI controller.

Use the mentioned four controllers respectively to control a multiple-input and multiple-output (MIMO) coupling system:

$$\begin{aligned} &G11=\frac{-5.72}{56s+1},G12=\frac{8e^{-81s}}{89s+1};\ G21=\frac{-2.80}{32s+1},G22=\frac{12e^{-90s}}{18s+1}\\ &G31=\frac{-4.80}{60s+1},G32=\frac{7e^{-70s}}{70s+1};\ G41=\frac{-3}{16s+1},G42=\frac{6e^{-60s}}{90s+1}. \end{aligned} \tag{107.18}$$

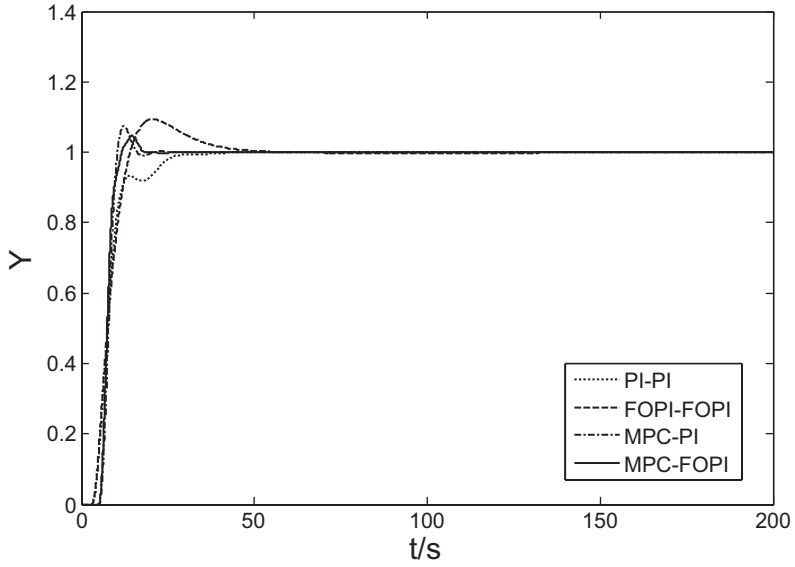


Fig. 107.4 Control effect of the single-cascade system without disturbance

G11 is a function that means input 1 to output 1, and G12 means input 1 to input 2 and so on. This system is a coupling system. It could be decoupled to several cascade systems, then the designed controller can deal with it.

The set point is $y_1=8$; $y_2=6$. The optimal FOPI-FOPI controller is

$$\begin{aligned} \mathrm{PI}_1^{-\lambda} &= 0.0327+0.234\mathrm{s}^{\,0.2607} \\ \mathrm{PI}_2^{-\lambda} &= 0.0113+0.272\mathrm{s}^{\,0.1912}. \end{aligned} \tag{107.19}$$

In this situation (Fig. 107.5), the *PI-PI* controller has no control effect. The performance index is, respectively,

$$J_{\text{MPC-FOPI}} = 118.0697; J_{\text{MPC-PI}} = 158.8055; J_{\text{FOPI-FOPI}} = 2132.7304. \tag{107.20}$$

Simulate (in Fig. 107.6) the four mentioned controllers to control the single-cascade system above under the measurable disturbance of the white noise (the mean power is 0.3 and the period is 3 s) and a measurable disturbance of the step ($r=0.1$).

It is shown that the MPC controller can effectively deal with the large-delay system and unmeasured disturbance in Figs. 107.5 and 107.6. In Fig. 107.6, it obviously shows that the FOPI controller features faster response ability and stronger robustness than the traditional PI controller. The MPC-FOPI controller is feasible in dealing with the delay and interference links in the cascade system, and also in the MIMO system, which can be decoupled to the cascade system.

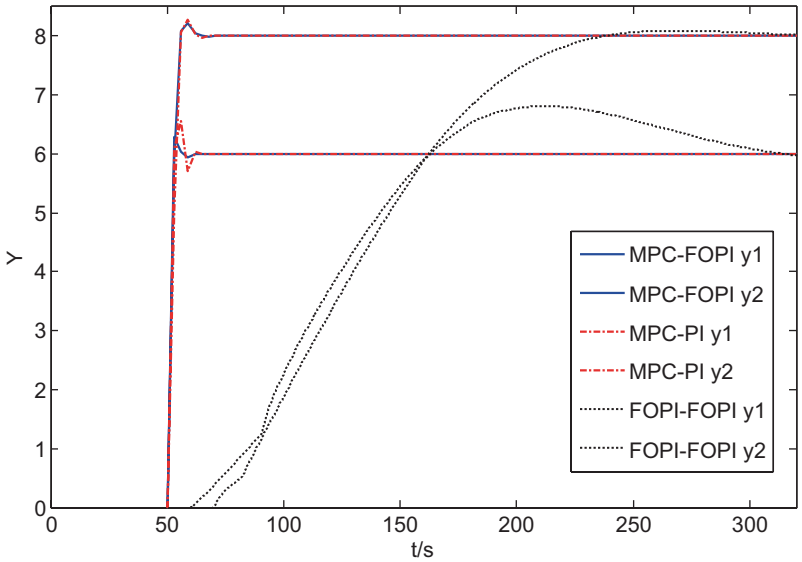


Fig. 107.5 A coupling system without disturbance

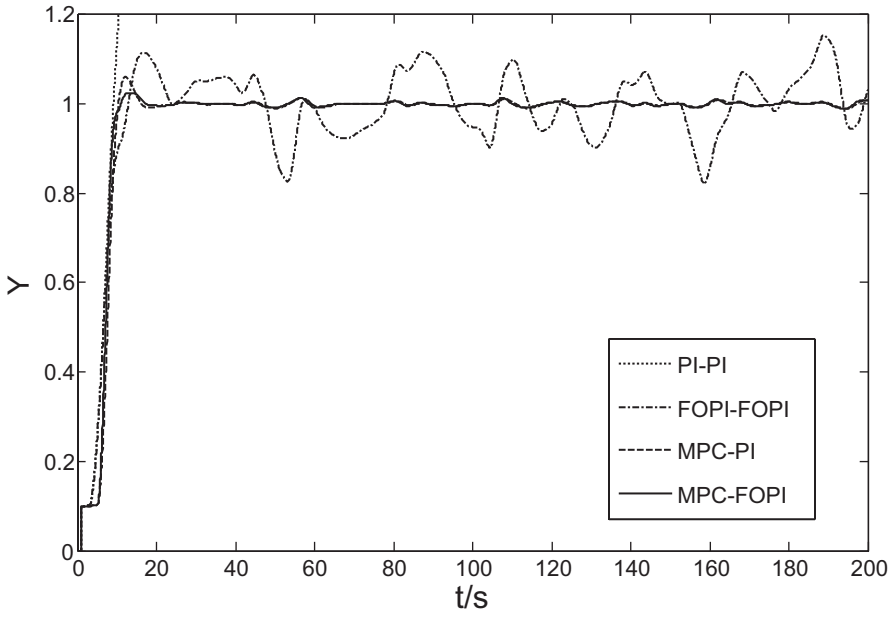


Fig. 107.6 A cascade system with disturbance

107.5 Conclusion

An MPC controller can deal with the pure delay system directly and has a good tracking performance. It also has a strong robustness while the FOPI controller has a faster response and disturbance rejection than the traditional PI controller. A new MPC-FOPI controller, which inherits the advantages of the MPC controller and the FOPI controller, is proved that it can better deal with the delay and interference links in the cascade system than the traditional PI-PI controller or MPC-PI controller.

The simulation results reveal that the MPC-FOPI controller has better control performance in dealing with the time-delay system, tracking the given value and rejecting the measured and unmeasured disturbance in the single-cascade and the coupling system as well.

References

1. Xi YG, Li DW, Lin S. Model predictive control—status and challenges. Beijing Acta Automatica Sinica. 2013;39(3):222–36 (In Chinese).
2. Hai L, Feng QP. The application of cascade PID control in UAV attitude control. Beijing Microcomput Inf. 2009;22(1):04–5 (In Chinese).
3. Flynn ME, O'Malley MJ. A drum boiler model for long term power system dynamic simulation. Power Syst IEEE Trans. 1999;14(1):209–17.
4. Xu M, Li S, Cai W. Cascade generalized predictive control strategy for boiler drum level. ISA Trans. 2005;44(3):399–411.
5. Singh R, Ierapetritou M, Ramachandran R. System-wide hybrid MPC-PID control of a continuous pharmaceutical tablet manufacturing process via direct compaction. Eur J Pharm Biopharm. 2013;85(3):1164–82.
6. Liu H, Yoo CK. Performance assessment of cascade controllers for nitrate control in a wastewater treatment process. Korean J Chem Eng. 2011;28(3):667–73.
7. Zhou R, Lou Z. Simulated research on DMC-PID cascade level control for dual vessel water tank. Shanghai Process Autom Instrum. 2011;32(10):63–5 (In Chinese).
8. Sen M, Singh R, Ramachandran R. A hybrid MPC-PID control system design for the continuous purification and processing of active pharmaceutical ingredients. Processes. 2014;2(2):392–418.
9. Guo W, Wang L, Zhou WP. Improved fractional-order PID model algorithmic control algorithm. Shenyang Control Eng China. 2011;5(1):022–3.
10. Podlubny I. Fractional-order systems and PI/sup/spl lambda//D/sup/spl mu//-controllers. Autom Control IEEE Trans. 1999;44(1):208–14.
11. Zhao C, Xue D, Chen YQ. A fractional order PID tuning algorithm for a class of fractional order plants. Mechatronics and Automation, 2005 IEEE International Conference, IEEE; 2005. p. 216–21.

Part IX
Computer Technology

Chapter 108
Functional Magnetic Resonance Imaging Based on Large Data

Weiyi Zeng, Sanyong Zou and Hao Zuo

Abstract Functional magnetic resonance imaging (FMRI) has been widely used in the field of medicine and physiology, which realizes neuro imaging without damage. FMRI was initially applied to the analysis of FMRI images; however, its statistical analysis was limited to the elementary processes because of the complexities of the FMRI images and the difficulties to establish the neuromotor time–space model. In this chapter, we focus on the analysis of FMRI images based on large data. Firstly, we make the time-correlation analysis of the data to reduce the size of the data. Then, we further investigate the spatial characteristics of valid signals, and compare the predicted signal and the original signal in the time domain and frequency domain. We find that there is a strong relevance of FMRI data in both time and space area, which indicates that the stimulation signal of the brain radiates to the surroundings of the stimulated point, and the signal is continuous in time, not an impulse.

Keywords Big data · Correlation analysis · FMRI images · Space domain · Time domain

108.1 Introduction

The neuroimaging technique has important significance for humans to achieve the ultimate scientific dream—"to uncover the mysteries of the brain". Functional magnetic resonance imaging (FMRI) is used in the field of medicine and physiology to

W. Zeng (✉)
Huazhong University of Science and Technology, 430074 Wuhan, China
e-mail: 15527302483@163.com

S. Zou
Tinghua University, 100084 Beijing, China
e-mail: zousy11@163.com

H. Zuo
Beijing Institute of Technology, 100081 Beijing, China
e-mail: 1120110342@bit.edu.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_108

make neuroimaging without damage become reality [1, 2]. Through the analysis of imaging data, we can get the information how human brain's neural state changes under different external stimulus conditions; thus, we can thoroughly understand the inner mechanism how human behavior is controlled by the brain and reveal the laws of brain nerves' movement. FMRI emerged in the 1990s, and the technology mainly focuses on the changes of the blood flow in various parts of the brain. FMRI uses the coupling relationship between blood flow conditions and nerve behavior to study the movement law of the brain [3]. Friston divides the research of brain function into two categories: one is used to analyze the space connection, which is called functional connection, and the other one is for the analysis of the time connection, which is called effective connection. They both analyze the interactions of neurons. There are two methods of the functional connection: one method is hypothesis driven, including correlation analysis, coherence analysis, and the generalized linear model (GLM); the other method is data driven, including independent component analysis, principal component analysis, and principal component analysis (PCA) cluster analysis. The methods of effective connection are divided into two categories according to whether the time information is included. Structural equation modeling (SEM), multivariate auto regressive modeling (MAR), and dynamic casual analysis (DCM) do not contain the time information; however, the Granger causality analysis contains the time information. Although statistical methods have been applied to the analysis of FMRI images in the early development of FMRI [4], due to the complexity of FMRI images, the neuromotor time–space model is difficult to be established, so its statistical analysis has been limited to a primary processing. Especially, there has not been substantive progress for the functional connection. In this chapter, we use the correlation analysis of the functional connection and the causality analysis of effective analysis [5, 6]. That the data is too large is the most serious problem in the analysis process; therefore, in this chapter, we first reduce the data size by making time-correlation analysis on data. Then, we further study the spatial characteristics of a valid signal, and compare the original signal and the predicted signal in time domain and frequency domain.

108.2 Methods

- Data preprocessing
  In this experiment, the method to reduce the data size is based on the characteristics of the problems and the following assumptions: we suppose the simulation signal in special region to be a pulse,and the signal generated by brain cannot be step. Therefore When a stimulus occurs, the impact will last for a long time, that is to say, the signal has a strong temporal correlation.
  Based on the above assumptions, the method I apply to reduce the data scale is the time-correlation analysis: For each spatial signal $S(_t)$, I take a time lapse and do an autocorrelation. The higher the correlation coefficient is, the closer the signal in this point is to the valid information; otherwise, the signal is approach-

ing the noise. It can be looking at the vast amount of data to a valid signal and follow-up study. Time domain waveform and frequency domain waveform obtain different autocorrelation maximum value by choosing different time lapses. Thus, we can find effective signal in a huge number of data and do subsequent research.

- Spatial characteristics of valid signals
  We assume that the signal also has a strong correlation with time. Work units of the brain are divided by the region; thus, the stimulating signal will be spread in space. Our conjectures are verified in the following experiments.
- The signal prediction
  We observe the signal curve in the region where the correlation coefficient is relatively large, and we speculate that the model is a bell-shaped pulse (Gaussian pulse). The universal expression of a Gaussian pulse is:

$$f(t) = E \times e^{-\left(\frac{t-t_0}{\tau}\right)^2}. \tag{108.1}$$

In this model, we suppose E to be the maximum value of effective signal whose correction value is the max one, and amended to 0.95E. t_0 is the time delay, and its value is 0 in the point with the maximum correlation coefficient. τ is the variance, and its value can be calculated:

$$\min(f(t)) = E \times e^{-\left(\frac{T-t_0}{\tau}\right)^2}. \tag{108.2}$$

T is the signal length; we can develop:

$$\tau = \frac{T}{\sqrt{ln\frac{E}{\min(f(t))}}}. \tag{108.3}$$

108.3 Experiments

- Data preprocessing
  Based on the above methods of the data preprocessing, we can get the time domain waveform and frequency domain waveform obtained from the signal of a time autocorrelation maximum value, which are shown in Fig. 108.1. Time domain waveforms of the four randomly gotten signals are shown in Fig. 108.2. The frequency domain waveforms of the corresponding four random signals are shown in Fig. 108.3.

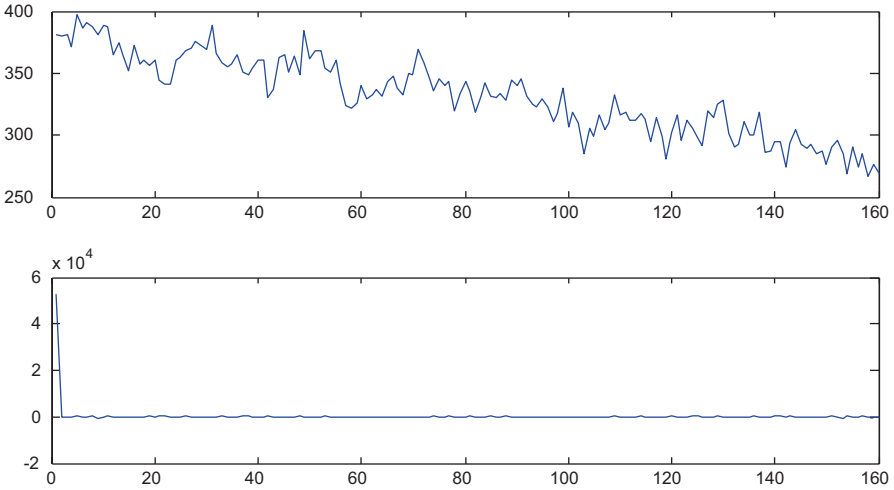


Fig. 108.1 Time domain waveform and frequency domain waveform obtained from the signal of a time autocorrelation maximum value

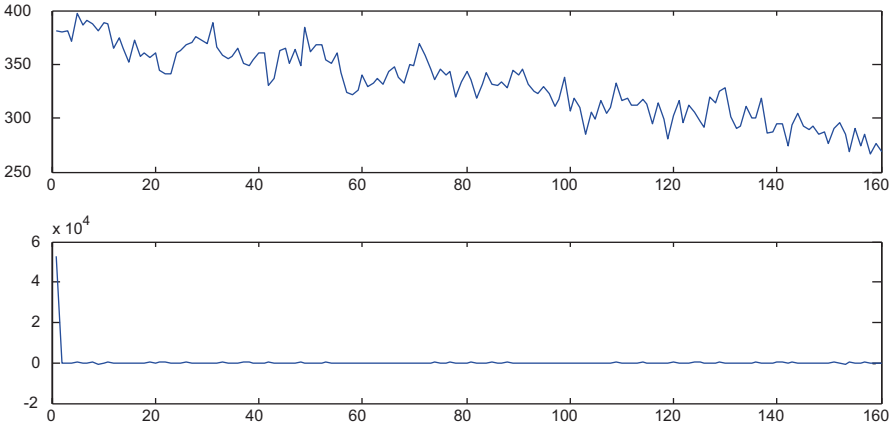


Fig. 108.2 Time domain waveforms of four randomly gotten signals

The results above show that:

- The obtained effective signal is consistent with assumptions. It is a deamplification in the time domain, and it has a high fundamental component and a low high-frequency component in frequency domain.
- Comparing the effective signal to the random signal, the random signals are similar to white Gaussian noise, and the high-frequency components is relatively low when compared to the low-frequency components.
  So, the experimental results are consistent with the expected results.
- Spatial characteristics of valid signals

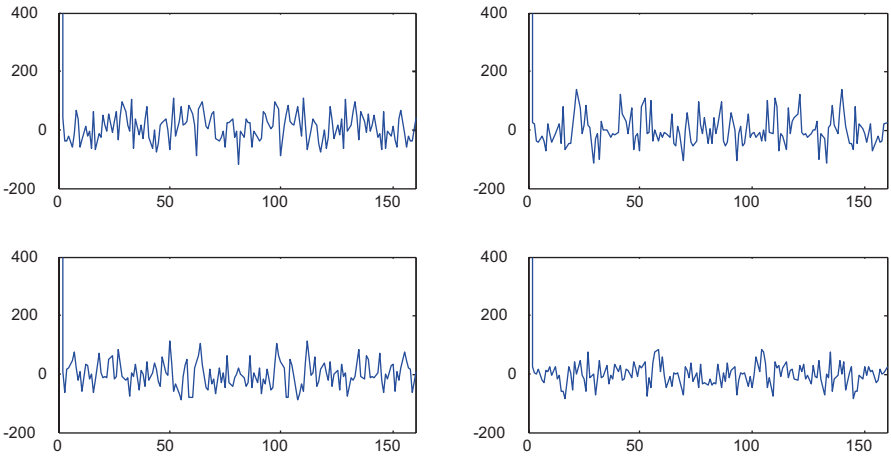


Fig. 108.3 The frequency domain waveform of the corresponding four random signals

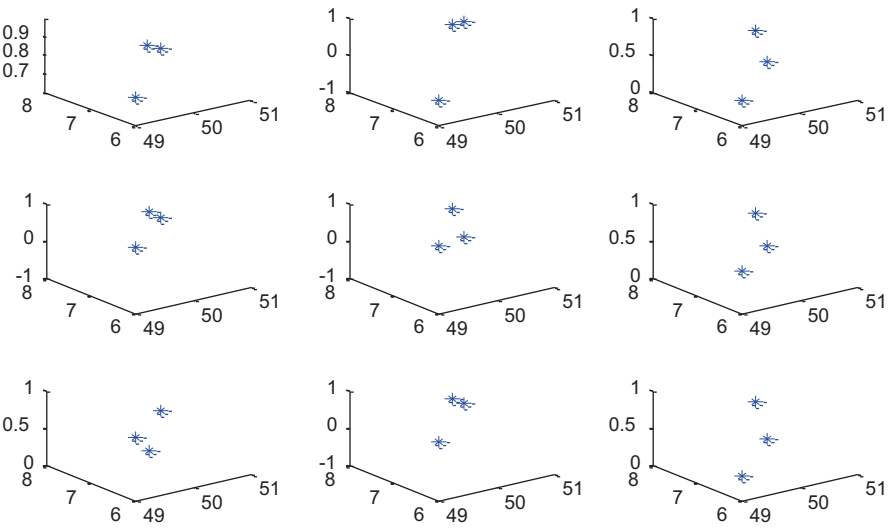


Fig. 108.4 A signal near the stimulation point and the autocorrelation coefficients of that signal

Firstly, we analyze the spatial correlation of the above stimulation points. Figure 108.4 shows a signal near the stimulation point and the autocorrelation coefficients of that signal.

Through the above-simulated results, we can find that the signal near the stimulated point does have a strong spatial correlation, in accordance with our original conjecture.

- The signal prediction
  We ensure that the correlation coefficient calculated is more than 0.9. The simulation model is shown in the following experiments (Fig. 108.5).

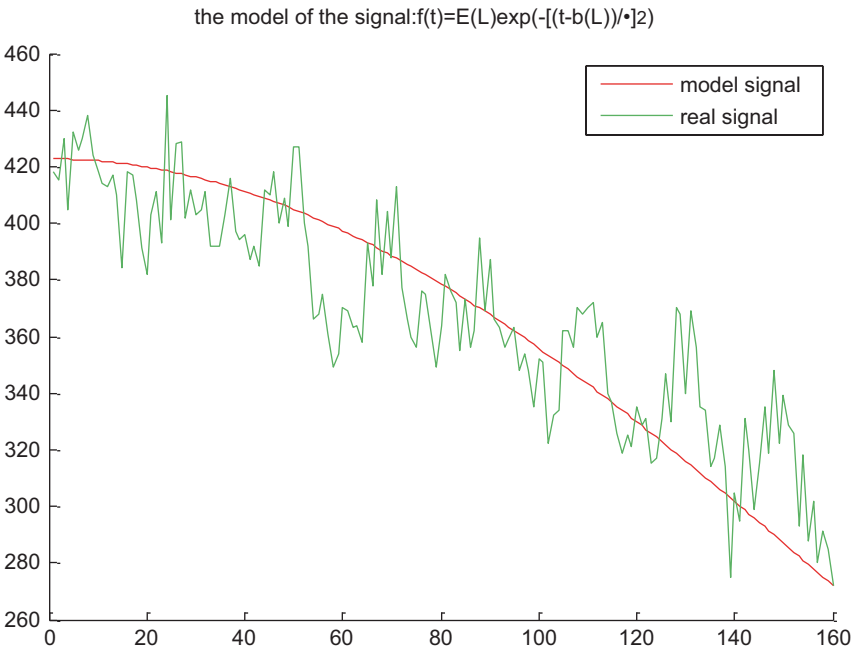


Fig. 108.5 The model of the signal: $f(t) = E(L)e^{-\left[\frac{(t-b(L))}{\tau}\right]^2}$

Taking the noise into account, we take the signal with the smallest correlation value to be the noise signal. Calculating again, we find that the correlation coefficient decreased instead of increasing. The simulation result is shown in Fig. 108.6.

When $t_0 \neq 0$:

Firstly, we get the propagation path according to the waveform of the signal by observing signals near the effective signal. Then, we take a point in the propagation path, and we choose the appropriate delay to simulate (Figs. 108.7 and 108.8).

Thus, we can conclude that the model of the signal transmitted by the brain after the magnetic resonance imaging (MRI) is the Gaussian pulse signal, and the signal spreads inside of the brain in a certain path rather than by omnidirectional spread. This conclusion has been reflected in the above correlation value point. In the field of biology, many values are Gaussian distribution, such as the amount of hemoglobin and the number of red blood cells. This may contain some of the essence of the organism.

108.4 Conclusion

The brain stimulation signal in the time domain is an approximated certain exponential decay time shift signal, and the signal in the frequency domain is consistent with the Fourier transform of the exponential signal. Therefore, a simple model

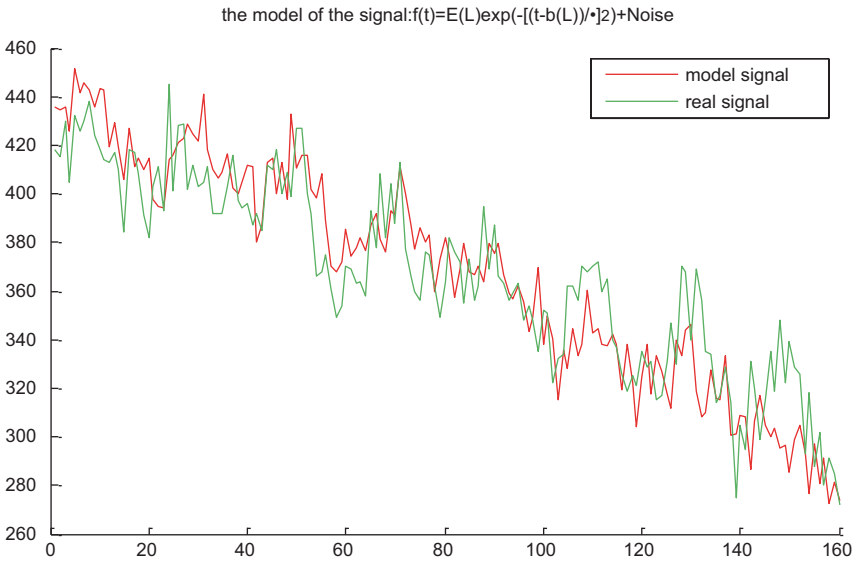


Fig. 108.6 The model of the signal $f(t) = E(L)e^{-\left[\frac{(t-b(L))}{\tau}\right]^2} + \text{Noise}$

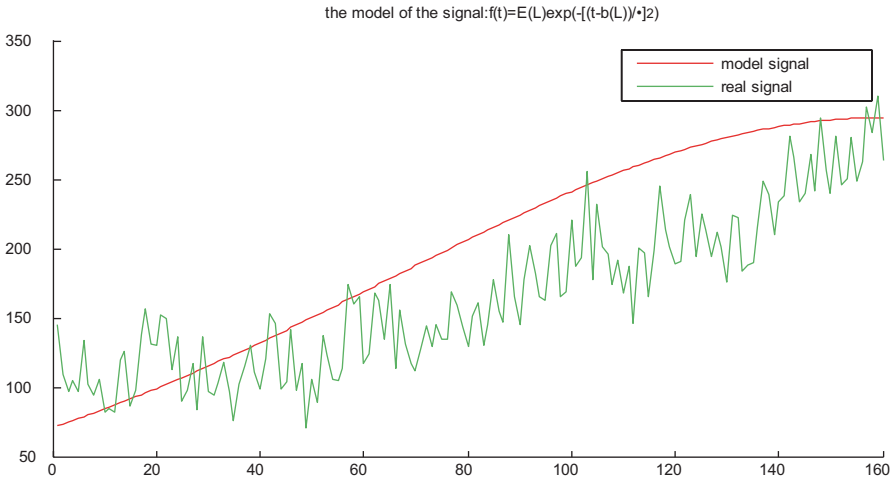


Fig. 108.7 Simulation without noise. $f(t) = E(L)e^{-\left[\frac{(t-b(L))}{\tau}\right]^2}$

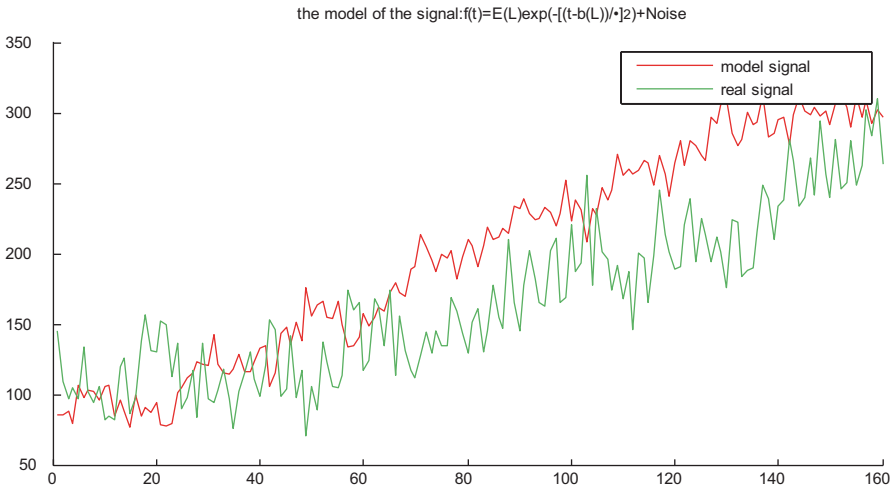


Fig. 108.8 Simulation with noise. $f(t) = E(L)e^{-\left[\frac{(t-b(L))}{\tau}\right]^2} + \text{Noise}$

for the brain stimulation signal is a time-shifted exponential signal, and the time correlation of the signal is high, which is also consistent with the results. Further, spatial-correlation analysis indicates that there is diffusion in the space of the signal in the simulated point, and this is called diffusion or transmission. Finally, when we make a simple signal-prediction based on the exponential model, we find that there is a small error between the predicted results and the actual results.

References

1. Hunter JV, Wilde EA, Tong KA, et al. Emerging imaging tools for use with traumatic brain injury research. J Neurotrauma. 2012;29(4):654–71.
2. Power JD, Barnes KA, Snyder AZ, et al. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage. 2012;59(3):2142–54.
3. Rose S. The Need for a Critical Neuroscience[J]. Critical neuroscience:A handbook of the social and cultural contexts of neuroscience. Oxford: Blackwell; 2012. 6(10):53–66.
4. Wolz R, Julkunen V, Koikkalainen J, et al. Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. PloS ONE. 2011;6(10):e 25446:3–5.
5. Vaitl D, Birbaumer N, Gruzelier J, et al. Psychobiology of altered states of consciousness. Psychol Conscious Theory Res Pract. 2013;1(1):2–47.
6. Vaitl D, Birbaumer N, Gruzelier J, Jamieson GA, Kotchoubey B, Kübler A, Strehl U, Lehmann D, Miltner WHR, Weiss T, Ott U, Sammer G, Pütz P, Wackermann J, Strauch I. Psychobiology of altered states of consciousness. Psychol Conscious Theory Res Practice. 2013;1(1):2–47.

Chapter 109
The Routing Protocol Based on the Aggregation of Routing Table in Cognitive Ad Hoc Networks

Tian Wang and Runhe Qiu

Abstract In cognitive ad hoc network, how to effectively use the limited spectrum resources to improve the performance of wireless networks has become a hot topic in recent years. We present a new routing protocol based on AODV (ad hoc on-demand distance vector) routing protocol-ORT-AODV in cognitive ad hoc networks. Firstly, the routing table and the total delay calculations are introduced for the secondary user who needs to switch between bands frequently in cognitive ad hoc network. A particular emphasis is given to routing table pooling mechanism, which adds route information collected from different channel to the routing table. Moreover, intermediate nodes initiatively updating routing table mechanism is proposed to tell the source node that the route is available or not. Finally, the respective simulation results demonstrate that it could comply with the requirements of cognitive ad hoc networks, effectively improve the system throughput performance and decrease transmission delay simultaneously.

Keywords Cognitive network · Ad Hoc · AODV · Delay · Routing protocol

109.1 Introduction

In recent years, cognitive radio is regarded as the technology to maximize the use of increasingly scarce radio spectrum resources. It was first proposed by Dr. Joseph Mitola [1] and considered to be the technology that can combine with ad hoc network. Although the traditional ad hoc network has many mature routing protocols, but working band in the cognitive ad hoc network changed over time and space. Therefore, routing protocol in the cognitive ad hoc network is a very important topic in wireless communication area.

T. Wang (✉) · R. Qiu
College of Information Sciences and Technology, Engineering Research Center of Digitized Textile & Fashion Technology, Ministry of Education, Donghua University, 200000 Shanghai, China
e-mail: wangtian4837@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_109

The studies proposed by international experts on routing protocols in cognitive ad hoc network can be summarized as follows: modify the existing ad hoc routing protocols, modify the MAC layer and design a cross-layer routing protocols, combination of both. Routing protocols are all proposed by modifying the existing routing protocols [2–4]. Routing protocol introduced is the cross-layer routing protocol [5]. Protocols are the combination of revising the existing ad hoc routing protocols and modify MAC layer [6–10]. AODV-cog is the routing protocol for cognitive radio CSMA/CA ad hoc network that allows node in the network has one transceiver for data transmission and one transceiver for control message exchange [11]. AODV-cog can meet cognitive radio demand in cognitive ad hoc network, but did not consider whether an available route existed before starting route discovery process.

All of the routing protocols mentioned above did not consider gathering routing information collected from different bands to find available route information before starting route discovery process when looking for available band.

The proposed ORT-AODV (overlay routing table-AODV) perceived available bands of the node through the MAC layer, added node routing information in different bands to the routing table and proposed initiatively updating routing table mechanism when the intermediate node is unavailable. The "ORT-AODV" is an abbreviation of overlay routing table-AODV, which means collecting routing information in different bands and adding to the routing table. This method can effectively improve the system throughput performance and decrease transmission delay.

109.2 Routing Problem Description in Cognitive Ad Hoc Network

109.2.1 Routing Table Update Problems

For ad hoc network, nodes work in a fixed band and update routing table in local, but in cognitive ad hoc network, nodes require switching between bands frequently for working band is dynamic over time and space. As Fig. 109.1 show, A is the source node, C is the destination node, A→B→C is an available path in the band 1, when switching to band 2, node A needs to restart routing discovery process to find a right path to C, A→B→C will be deleted once starting route maintenance mechanism for the route is unavailable, if A backs to the band 1, A needs to restart the discovery process instead of A→B→C, which will increase delay.

109.2.2 Intermediate Node Switching Band Problem

When intermediate nodes switch to the other band, routing recovery process will be started to find an available route to the destination node. If no route is found, the current route will be considered as invalid, and the source node needs to start

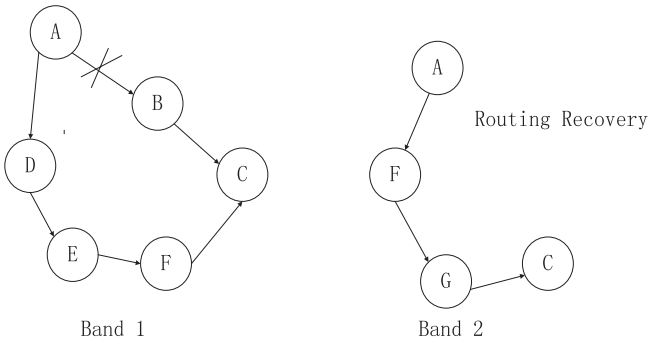


Fig. 109.1 Restarting routing discovery mechanism when switching to the other band

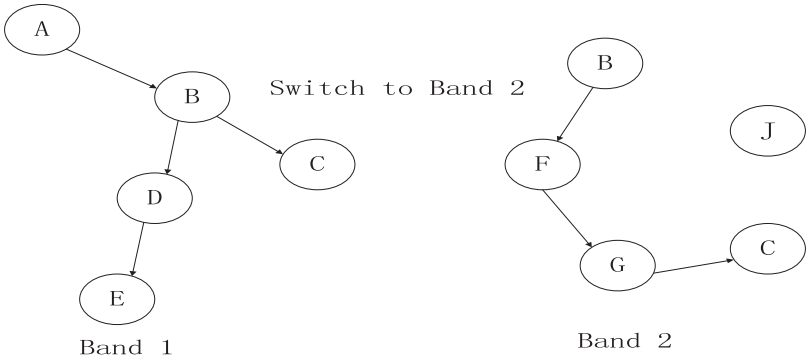


Fig. 109.2 Intermediate node switching between bands

routing discovery process, which will cause data loss and latency problem. Routing recovery process is shown in Fig. 109.2, node A is the source node, node C is the destination node, when node B switch to band 2 by the primary user interference, B→C is an unavailable route, node B starts routing recovery process. If a valid route is found, like B→F→G→C, the data will transmit via A→B→F→G→C. Otherwise, node A will start routing discovery process.

109.3 The Proposed Cross-layer Routing Algorithm—ORT-AODV

To solve the above questions, the paper proposed ORT-AODV routing protocol, which modified the AODV routing table, added the routing table pooling mechanism and the intermediate nodes initiatively updating routing table mechanism to AODV routing protocol. ORT-AODV can effectively improve the system throughput performance and decrease transmission delay.

Table 109.1 Routing table

Did	Dseqo	NextHop	RouteCast	Flag	LifeTime	Reserve
		Currrent_Freq_Band	Next_SOP	Flag_Switch		Next_Freq_Band

109.3.1 The Design of Routing Table

In the traditional AODV routing protocol, the routing table contains the latest serial number of the destination node, Next-hop address, flag, routing cast, lifetime, and so on. But the routing table in traditional AODV does not meet the characteristics of the cognitive ad hoc network, and you need to expand the routing table below (Table. 109.1).

Node current working band (Current_Freq_Band): shows the band that is in use. Current_Freq_Band can also be regarded as the index of route information in different bands, which is import in the routing table pooling mechanism.

Next-hop available spectrum (Next_SOP): stores available spectrum for the next hop. If the route is not used within the expiration time, or the next-hop node is regarded as unreachable, Next_SOP will offer available spectrum to the next node to do routing recovery process or routing discovery process.

Next hop node working band (Next_Freq_Band): shows the working band of the next node.

Flag_Switch: shows the next node is reachable or not. Default value is 1, which means that the next node is reachable. Otherwise, the value of Flag_Switch was changed by starting the initiatively updating routing table mechanism on next node.

109.3.2 The Routing Table Pooling Mechanism

Since nodes need to frequently switch between bands, the routing discovery mechanism is started to find an available route, will cause latency problem. The routing table pooling mechanism is proposed to reduce latency, which adds route information collected from different channel to the routing table index in the current working band. If the nodes switch to the other bands, get route information from the routing table by the current working band, which will avoid starting routing discovery again when re-switch to ever used band. Working process is shown in Fig. 109.3.

109.3.3 The Initiatively Updating Routing Table Mechanism

The initiatively updating routing table mechanism is proposed to solve latency problem caused by intermediate node switching to other bands. Working process is shown in Fig. 109.4.The intermediate node will check the available route to

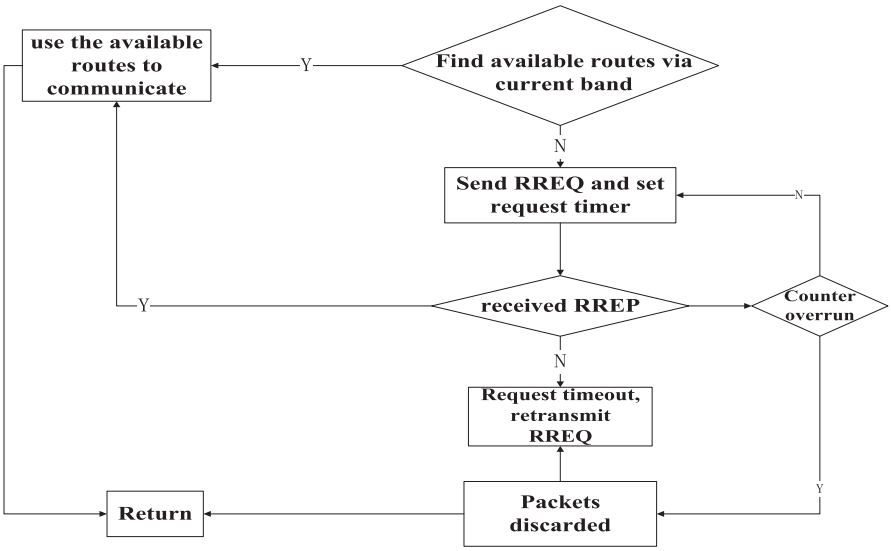


Fig. 109.3 Working process for the routing table pooling mechanism

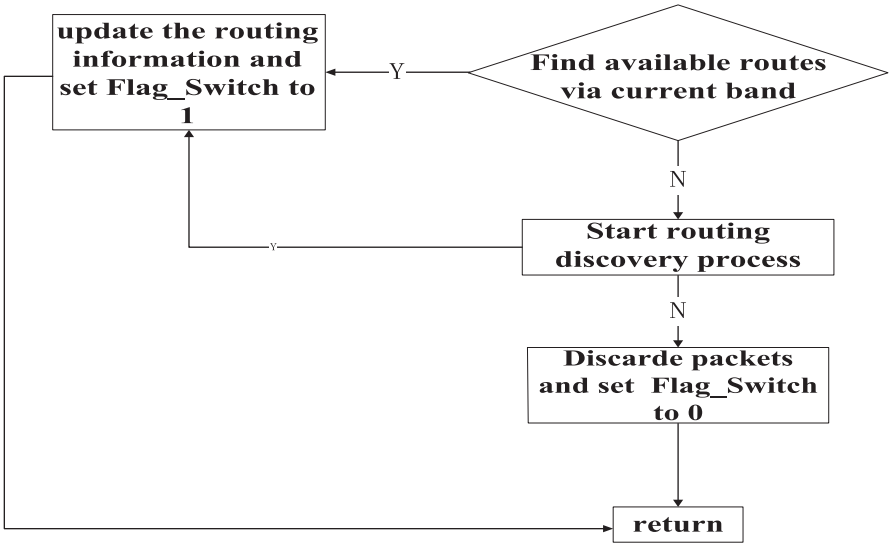


Fig. 109.4 working process for the initiatively updating routing table mechanism

the destination node firstly in the routing table by the current working band when switching frequencies. If there exists an available route, set Flag_Switch to 1 in the routing table of source node. Otherwise, start the routing discovery process to find an available route to the destination node. If no such route is found, delete the route from the routing table and set Flag_Switch to 0 in the routing table of source node.

If suppose time to transmit RREQ and RREP packet is same, let the time be h, the total delay in total delay can be expressed as

$$\text{Dtotaldelay} = 2 * m * h + Dtotal(a,b) \tag{109.1}$$

Where sw_delay means switching, m is hop.

109.4 Simulation and Analysis

NS2 is used to simulate ORT-AODV routing protocol. We compare CR-AODV routing protocol with ORT-AODV routing protocol in throughput and delay. CR-AODV routing protocol is the improvement of AODV routing protocol in cognitive ad hoc network, which can efficiently use and exploit the multiple available channels and interfaces through a competent assignment strategy and an adequate route metric. Parameters used in the emulation shown in Table 109.2.

Figures 109.5 and 109.6 show the average throughput and delay changed by numbers of nodes and primary users. ORT-AODV-X represents the ORT-AODV protocol performance in the case of X numbers primary user, CR-AODV-X represents the CR-AODV protocol performance in the case of X numbers primary user. According to Fig. 109.5, the network throughput using the above-mentioned protocols increased as the number of nodes increased, decreased with the increase of primary users, but the throughput using ORT-AODV protocol is greater than using CR-AODV protocol. Average delay shown in Fig. 109.6 decreased as the increase of nodes numbers, but the improvement on average delay by using ORT-AODV protocol is more obvious.

Table 109.2 Simulation parameters

Type	CBR
Packet size	512 bytes
Coverage area	800 m*800 m
Transmission	660 mW
Received power	400 mW
Simulation time	120 s
Node maximum speed	10 m/s

109.5 Conclusion

Based on the AODV protocol, this paper proposes the ORT-AODV routing protocol in cognitive ad hoc network. It collects available bands through the MAC layer, adds the route information found in different frequency bands to the rout-

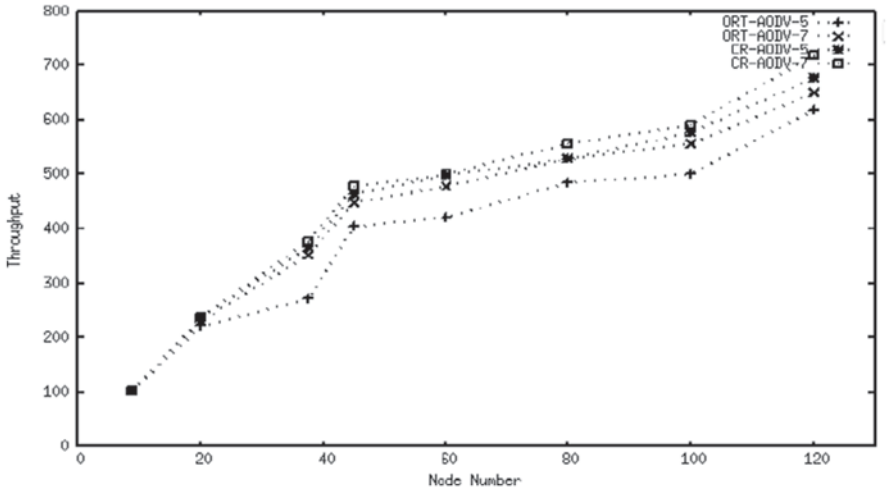


Fig. 109.5 Throughput changed by node numbers

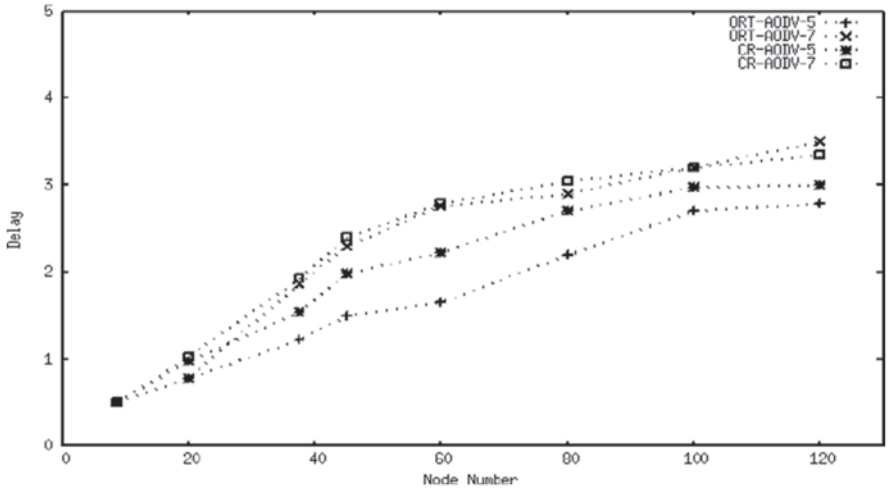


Fig. 109.6 Delay changed by node number

ing table and starts the initiatively updating routing table mechanism when the intermediate node switching to other bands. The proposed ORT-AODV protocol can effectively reduce the network delay and reduce the time to discover the route when switching to other channels. Through simulation, we found that ORT-AODV protocol can effectively improve the system throughput and reduce delay. But the initiatively updating routing table mechanism is based on certain additional delay at the expense. The focus of our future research is how to further reduce the delay of cognitive ad hoc networks, and balance the performance of high throughput and low delay.

References

1. Mitola J, Maguire GQ. Cognitive radio: making software radios more personal [J]. Personal Commun IEEE. 1999;6(4):13–8.
2. Rehman RA, Sher M, Afzal MK. Efficient delay and energy based routing in cognitive radio ad hoc networks[C]//Emerging Technologies (ICET), 2012 International Conference on. IEEE; 2012: 1–5.
3. Tang F, Barolli L, Li J. A joint design for distributed stable routing and channel assignment over multi-hop and multi-flow mobile ad hoc cognitive networks[J]. Ind Inform IEEE. 2012;10(2):1606–15
4. Cacciapuoti AS, Caleffi M, Paura L. Reactive routing for mobile cognitive radio ad hoc networks[J]. Ad Hoc Netw. 2012;10(5): 803–15.
5. Shen J. Research for cross-layer routing protocol in cognitive networks[D]. Beijing: Beijing University of Posts and Telecommunications; 2010.
6. ZHENG Teng-fei, ZHU Qi. Cross-layer routing protocol for cognitive Ad Hoc networks[J]. J Comp Appl. 2011;31(A02):9–13.
7. Chowdhury KR, Akyildiz IF. CRP: a routing protocol for cognitive radio ad hoc networks[J]. IEEE J Sel Areas Commun. 2011;29(4):794–804.
8. Chi S. Design and Implementation on the Cognition-based Routing Protocol of the Ad Hoc[D]. Xi'an: Xidian University; 2011.
9. Abbagnale A, Cuomo F. Gymkhana: a connectivity-based routing scheme for cognitive radio ad hoc networks[C]//INFOCOM IEEE Conference on Computer Communications Workshops, 2010. IEEE, 2010: 1–5.
10. HU YJ, LUO T, SHEN JL. An improvement of the route discovery process in AODV for Ad Hoc network[C]//2010 International Conference on Communications and Mobile Computing. Washington, DC: IEEE Computer Society; 2010. Pp. 458–61.
11. Wu C, Ohzahata S, Kato T. A routing protocol for cognitive radio ad hoc networks giving consideration to future channel assignment [J]. Comput Netw (CANDAR). 2013;1(1):227–32.

Chapter 110
A High-Speed LVDS Driver Design in 0.35-μm CMOS Technology

Zhongyan Wang, Ai Guo and Yan Pu

Abstract With the development of new technologies, the operating frequencies on chip are increasing at a faster rate, such as computational methods, utilization of high-frequency clocks, digital circuits, etc. The process-technology-independent I/O standard, low-voltage differential signaling (LVDS), is basically developed for low-voltage, low-power, low-noise, and high-speed I/O interfaces. Low power is owing to the use of very small differential swing, while low noise is owing to essential nature of the differential circuits. Based upon ANSI TIA/EIA-644 LVDS standard, this paper presents a low-voltage and high-speed LVDS driver. A Common-mode feedback (CMFB) and a pull-up/down circuits were suggested as carried out by a standard 0.35-μm complementary metal oxide semiconductor (CMOS) process with a die area of 0.115 mm^2. The measured results present that the driver works well at 1.5 Gbps, and the static current is less than 11.5 mA under 3.3 V.

Keywords CMFB · High speed · LVDS · Rise/fall time

110.1 Introduction

LVDS driver plays an important role in point-to-point communication with more and more data stream quantity and increasing data rate, as microprocessor main board, high-speed ADC/DAC, optical transmission links, etc [1].While the scaled CMOS technology continues to enhance the speed of operating speed but the I/O interface still confines the rate of data process, so the LVDS I/O interface becomes necessary. The point-to-point LVDS interface is shown in Fig. 110.1. Saving by a different scheme for transmission, termination, and a low-voltage swing, LVDS completes significant power [2]. LVDS standards put relatively stringent requirement on the tolerance influencing the output levels, gaining interesting design issues if low-cost solutions with neither external components nor trimming procedures are required. In the meantime, the very large tolerance for the common-mode

Z. Wang (✉) · A. Guo · Y. Pu
Sichuan Institute of Solid-State Circuits, 400060 Chongqing, China
e-mail: zhongyan918@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_110

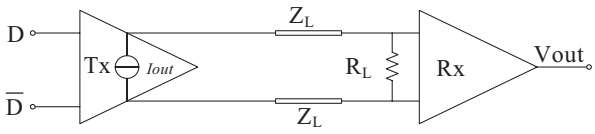


Fig. 110.1 Schematic of LVDS interface

voltage requires several design improvements by the CMFB circuit. Additionally, a pull-up/down circuit is integrant, while the LVDS operates at the rate of Gb/s or above for improving the rise/fall time.

This paper describes the design and the implementation of a low-voltage and high-speed LVDS driver in 0.35-μm CMOS technology. Section 110.2 discusses the current LVDS driver mode and design challenges for high-speed operation. In Sect. 110.3, a CMFB and a pull-up/down circuit is discussed, and the simulation results are given. Finally, the experimental results and conclusions are addressed.

110.2 LVDS Driver Circuit

110.2.1 Standard Requirement of the LVDS Driver

Table 110.1 briefly describes the main requirements of an output driver circuit based on the ANSITIA/EIA-644 LVDS standard. Regarding these requirements, it is very difficult to achieve the desired rise and fall time using a single 3.3V supply voltage. Therefore, a very careful design strategy is proposed to satisfy all these specification. Based on the LVDS standard, the V_{OD} should also be able to drive an external 100 Ω termination resistor [3].

110.2.2 Typical Driver Circuit

As shown in Fig. 110.1, a differential data transmission has been employed in a point-to-point LVDS interface. As a matter of fact, the load resistor at the receiver,

Table 110.1 The Main LVDS driver Requirements

Parameter	Min.	Max.	Unit
V_{OS} (Common-mode voltage)	1.125	1.375	V
V_{OD} (Differential voltage swing)	247	450	mV
ΔV_{OS} (Acceptable mismatch on V_{OS})	–	50	mV
ΔV_{OD} (Acceptable mismatch on V_{OD})	–	50	mV
t_R, t_F (rise and fall time)	0.26	1.2	ns

providing the optimum line impedance, is driving by a switch bridge in LVDS driver. Through the current or voltage mode, correct LVDS driver functionality can be achieved. In Junsheng Lv's paper [4], we can find the equivalent circuit diagram of the current mode and voltage mode LVDS driver.

Generally, the current mode drivers have better reflection performance due to the high-output impedances, which can minimize the impact output impedance on the chip. While the voltage mode drivers consume less power, the impedance of output often exhibits nonlinear behavior during the signal switching [5].

110.2.3 CMFB Circuit

Stabilization of the CM loop is difficult owing to the large size of output transistors. A high-frequency op-amp can build up an internal CMFB circuit to stabilize the CM signal in traditional way. But the result is not ideally due to the large size of the LVDS driver switches. So the CMFB op-amp should not only have high gain but also have high bandwidth [6].

110.3 Proposed Circuits

In order to improve the speed of data rate, a steep rise/fall time is necessary. But the parasitical capacitance and intrinsic performance of transistor confine the drive current. Increasing a charge/discharge circuit between the control signals of switches to produce short periodic pulses at the data's transition edges is a prevalent method to solve the issue, and the detailed expatiation is in Mingdeng Chen' s paper [1]. These current pulses are used to charge/discharge the parasitic capacitors and so pull-up/down the switchable current source gate voltages.

In this chapter, a novel pulling up/down circuit, using two auxiliary pulses to improve the ability of current source at the edges of switches changing, is proposed. The right part of Fig. 110.2 is the core of the LVDS driver, and the way to get a steep rise/fall time is illustrated in Fig. 110.3. The current switch, consisted by M1 ~ M4, is controlled by C and D. The driver outputs V_{OUTP} and V_{OUTN} are coupled to the transmission line. When D = 1, M3, and M2 are switched on, M1 and M4 are switched off. Thus, V_{OUTP} is pulled high and V_{OUTN} is pulled low; when D = 0, M2 and M3 are switched off, M1and M4 are switched on. Thus, V_{OUTP} is pulled low and V_{OUTN} is pulled high, generating logic "1" and"0" levels [7].

As shown in Fig. 110.3, a bandgap reference provides the reference voltage V_{REF}, reference current I_{REF} and bias voltage V_A and V_B, due to the strict limitations placed on the parameters of the driver by the standard.

The proposed CMFB amplifier consisted by M5 ~ M8 can get a tail source current produced by the current mirror array, the M14 ~ M17, and the M9 ~ M13. The magnitude is determined by the scale of the current mirror and source current I_{REF}. It is the same as the driver current I_{LOAD} produced by M_{PB1}, M_{PB2}, $M_{NA1,}$ and M_{NA2},

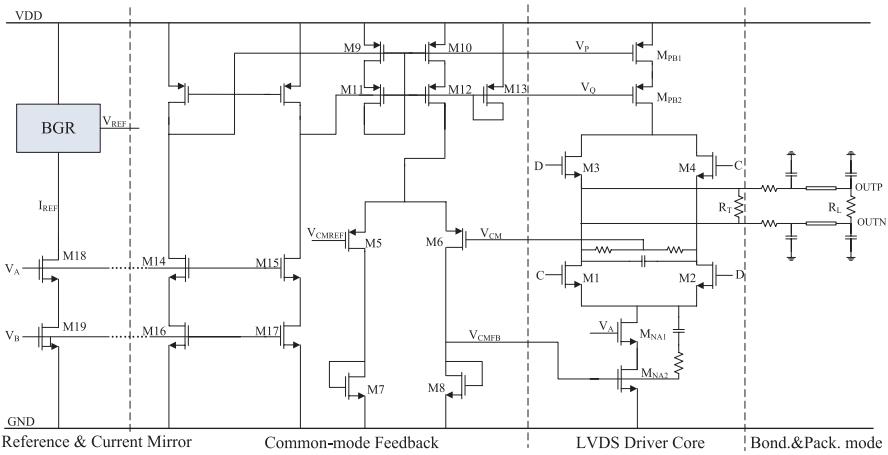


Fig. 110.2 CMFB and LVDS Driver Core circuit

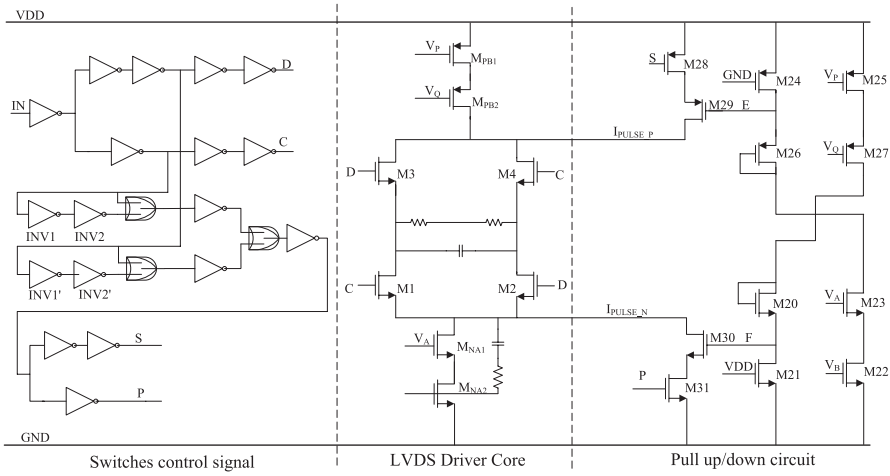


Fig. 110.3 Proposed Pull-up/down and Switch control circuit

which is on a large scale compared to the M18 and M19. Then, the V_{OD} of the driver is determined as

$$V_{OD} = I_{LOAD} \cdot R_T = N \cdot I_{REF} \cdot R_T \tag{110.1}$$

where N is the scale of the current mirror array.

The function of stable V_{CM} relies on regulating the I_{LOAD} finely by the output of amplifier, V_{CMFB}. Because of the requirement of sufficient drive ability and the capable responding to the current change, the large size of M5 and M6 is necessary.

The control signals C, D and S, P are complementary, produced by the logical circuit in Fig. 110.3, with a single-ended input signal. The detailed timing is shown in Fig. 110.4. Compared to the switch moment of D, the pulse signal S changed half a pulse width ahead, and the pulse width of S is determined by the size of INV1 and INV2 due to the NAND logical relation of nodes A and B. An equivalent circuit as shown in Fig. 110.5 is used to model the switch characteristic of the INV. The effective capacitance is given by

$$C_{IN} = (3/2) \bullet (C_{OX1} + C_{OX2}) = C_{INN} + C_{INP} \tag{110.2}$$

$$C_{OUT} = C_{OX1} + C_{OX2} = C_{OUIN} + C_{OUTP} \tag{110.3}$$

where$C_{OXI}$and $C_{OX2}$are the gate oxide capacitance of the NMOS and PMOS in INV, then the delay time of the INV1 and INV2 can be expressed as

$$t_{DL} = 0.7 \bullet (R_N + R_p) \bullet C_{TOT} \tag{110.4}$$

$$C_{TOL} = C_{IN} + C_{OUT} = (5/2) \bullet (C_{OX1} + C_{OX2}) \tag{110.5}$$

TheC_{OXI}, C_{OX2}, R_N, R_P are important parameters of the transistor, and delay time can obtain hands down from the W/L of the transistor.

The key of LVDS driver is the switches control and pull-up/down circuit. The rise/fall time is important for high speed of data rata, as shown in Fig. 110.6, if the rise/fall time takes up large scale of the period of the data rata, or the period of data rata and rise/fall are comparable with the speed up, the error code will happen, resulting a bad transmission. So, a pull-up/down circuit is necessary in Fig. 110.3, M24 and M26 are the loads of the NMOS current mirror consisted by M22, M23, to

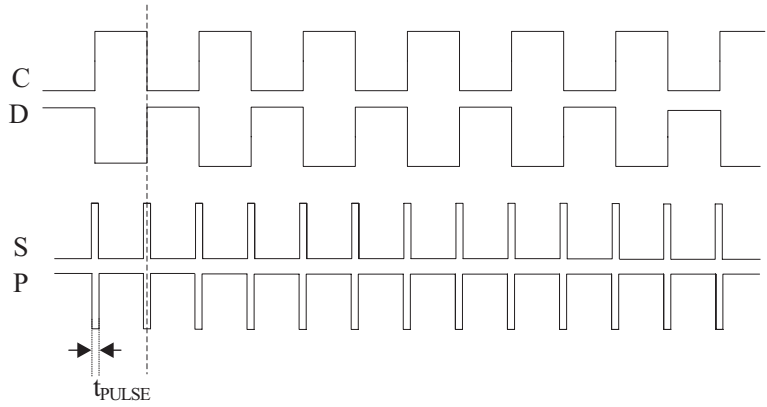


Fig. 110.4 Timing of Control signals of the Switches and Pull-up/down signals

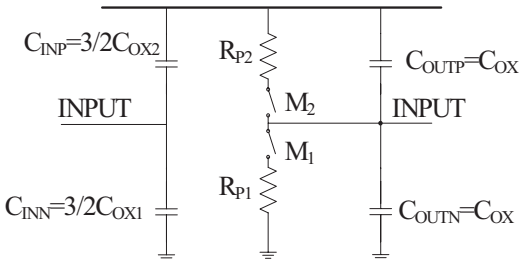


Fig. 110.5 Switch Characteristic of CMOS INV

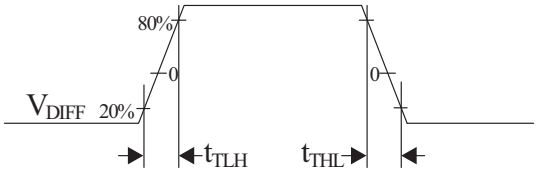


Fig. 110.6 Description of the rise/fall time

produce the bias voltage at node E. And M20 and M21 are the loads of the PMOS current mirror consisting of M25, M27, to produce the bias voltage at node F. With the proper bias voltage at node E and F, a very large saturation current flows through the M29 and M30 under the control of S and P. The current pulse $I_{PULSE\text{-}P}$ and $I_{PULSE\text{-}N}$ could be several times bigger than the drive current I_{LOAD} with a large size W/L of the M20 and M30, about 800/0.35. The width and amplitude of impulse are 200 ns and 10 mA, respectively.

110.4 Simulation and Measurement Results

The driver circuit was simulated using SMIC 0.35 μm CMOS technology. In order to make the circuit more close to the practical environment, the ESD, package, bond wire, and transmission line effects have been all considered in the circuit simulation. The load capacitance is 15 pF, and the load resistor is 100 Ω here.

In Fig. 110.7a, when the pull-up/down circuit is disabled, at the moment the gate voltage starts to change, the M1/M4 is not switched on along with the gate voltage quickly due to the lack of drain current, as the V_{GS} of M1/M4 is proportional to the square root of the drain current. As a result, the waveform of V_{OD} produced by the load current through the terminal resistor is shown as curve A. When the pull-up/down circuit is enabled, the switch M1/M4 get additional drain current to turn on while the gate voltage changed, and the V_{OD} is indicated by curve B. Obviously, the curve B has a cliffer rise/fall time. The proposed CMFB works well from -40℃ to 85℃, and the maximum output common-mode voltage deviation is only 3.2 mV, about 20.5 ppm/℃ as shown in Fig. 110.7b.

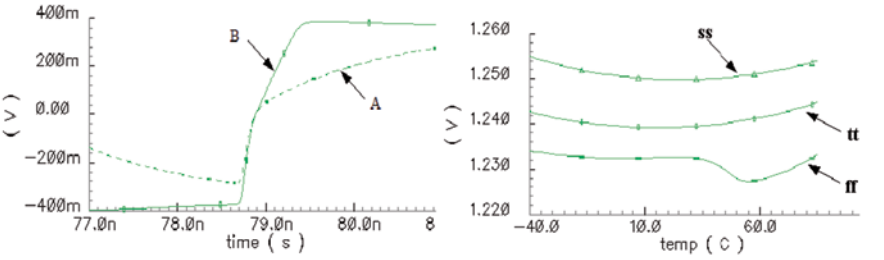


Fig. 110.7 **a** Simulated output differential signal **b** Simulated output common-mode signal

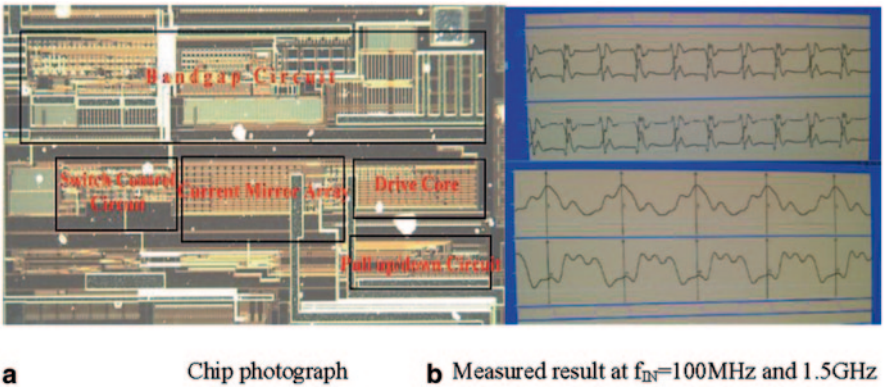


Fig. 110.8 **a** Chip photograph **b** Measured result at f_{IN} = 100 MHz and 1.5 GHz

The driver is carried out by a standard 0.35-μm CMOS technology, with a voltage supply of 3.3V. The die size is 0.115mm^2. The chip photograph is shown in Fig. 110.8a. The measurement presents that the driver can work well at a data rate of 1.5Gbps with 367 mV differential voltage swing. The measured results in Fig. 110.8b showed the rise/fall time decrease a lot because of the pull-up/down circuit, and the measured result is 260 ps.

The main parameters were compared with some previously reported work in Table 110.2. As the analysis before, the power consumption, die size, and data rate have advantages to other works [1, 8], especially the additional pull-up/down circuit reduces the rise/fall time markedly relative to the Vladimir Bratov' s paper[5]. The proposed pull-up/down circuits mainly contribute to the technology of LVDS driver in the chapter. Recently, the LVDS drivers achieve a higher data rate with an advanced process and technology. Hazem W. Marar' s paper lists a LVDS driver with a data rate of 3Gbps by adding a pre-driver in 0.18-μm CMOS process [9]. By using a positive-feedback and impedance matching technique in a 2.5V/1.2VSiGe BiCMOS process, the data rate of LVDS driver can achieve 10Gbps in Khaldoon Abugharbieh' s paper [10]. So, the driver in the chapter still has prodigious advancement in data rate and power consumption.

Table 110.2 LVDS driver performance comparison

Parameter	Ref. [1]	Ref. [5]	Ref. [8]	This paper
Process	0.35-μm CMOS	0.35-μm BiCMOS	0.35-μm CMOS	0.35-μm CMOS
Output voltage swing (mV)	340	300	412	368
Maximum data rate (Gbps)	1.2	1.5	1.2	1.5
Rise and Fall time (ps)	–	400	–	260
Power supply (V)	1.8	3.3	3.3	3.3
Power consumption (mW)	12.8	18.5	43	11.55
Die size (mm^2)	0.14	–	0.17	0.115

110.5 Conclusion

In this chapter, a high-speed LVDS driver has been developed. The proposed circuit is a pull-up/down circuit to improve the performance of rise/fall time. The chip was implemented by a 0.35-μm CMOS process. The driver operates with the data rate up to 1.5 Gbps, in accordance with other parameters shown in Table 110.2 with the LVDS standard with significant margin.

References

1. Chen M, et al. Low-voltage low-power LVDS driver. IEEE J Solid-State Circuits. 2005;40(2):472–9.
2. Boni A, Pierazzi A, Vecchi D. LVDS I/O interface for Gb/s-per-pin operation in 0.35-μm CMOS. IEEE J Solid-State Circuits. 2001;36(4):706–11.
3. Electrical characteristics of low-voltage differential-signal (LVDS). Interface circuits, TIA/EIA-644, National Semiconductor Corp., ANSI/TIA/EIA, 1996.
4. Lv J, Ju H, Yuan L, Zhao J, Zhang F, Wu B, Jiang J, Zhou Y. A high speed low jitter LVDS output driver for serial links. Analog Integra Circ Sig Process. 2011;68(3):387–95.
5. Bratov V, Binkley J, Katzman V, Choma J. Architecture and implementation of a low-power LVDS output buffer for high-speed applications. IEEE Trans Circuits Syst I Reg Papers. 2006;53(10):2101–8.
6. Tajalli A, Leblebici Y. A slew controlled LVDS output driver circuit in 0.18 μm CMOS technology. IEEE J Solid-State Circuit. 2009;44(2):538–48.
7. Gupta H, Paramar RM, Dave PK. High speed LVDS driver for SERDES. International Conference on Emerging Trends in Electronic and Photonic Device & Systems, 2009. ELECTRO' 09. International Conference on. IEEE;2009: 92–5.
8. Boni A, Pierazzi A, Vecchi D. LVDS I/O interface for Gb/s-per-pin operation in 0.35-μm CMOS. IEEE J Solid-State Circuits. 2001;36(4):706–11.
9. Marar HW, Abugharbieh K, Abdel-Karim Al-Tamimi. A Power Efficient 3-Gbits/s 1.8V PMOS-Based LVDS Output Driver," IEEE international conference on electronic, circuits and systems, IEEE, pp. 240–3, Dec. 2012
10. Abugharbieh K, et al. An ultralow-power 10-Gbits/s LVDS output driver. IEEE Trans Circuits Syst- I Reg Papers. 2010;57(1):262–9.

Chapter 111
Face Recognition Based on Cross-Validation by Support Vector Machine

Jingfang Hu and Lin You

Abstract Support vector machine (SVM) is a kind of machine learning based on statistical learning theory. It shows unique advantages in the small-sample, nonlinear, and high-dimension pattern recognition. Principal component analysis (PCA) is a multivariate analysis technology for feature extraction. In this chapter, we propose a new method of face recognition based on PCA and SVM. It applies PCA to extract face feature and uses SVM combined with cross-validation (CV) to classify face images. CV is a good method of parameter optimization in SVM. We conduct the recognition experiment on the Cambridge ORL database. Compared with other methods, the accuracy rate of face recognition is up to 89.5%. It is shown to be an effective method.

Keywords Face recognition · Support vector machine · Principal component analysis · Cross-validation

111.1 Introduction

Because of the development in demand of public security, judicial application, and identification, face recognition becomes a hotspot in pattern recognition. There is an obvious difference in identifying the same face matrix by computer, due to factors such as illumination, expression, shelter, and posture. Therefore, we should choose relatively stable characteristics when we extract facial features [1]. Principal component analysis (PCA) [2] is an effective method for face feature extraction that can retain the required feature vectors and realize the dimension reduction. Every face image in the database is represented as a vector of weights, which is the projection of the image to the basis in eigenface [3]. The extracted features would be

J. Hu (✉) · L. You
College of Communication Engineering, Hangzhou Dianzi University,
310018 Hangzhou, China
e-mail: hujingfang42@126.com

L. You
e-mail: linyou@hdu.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_111

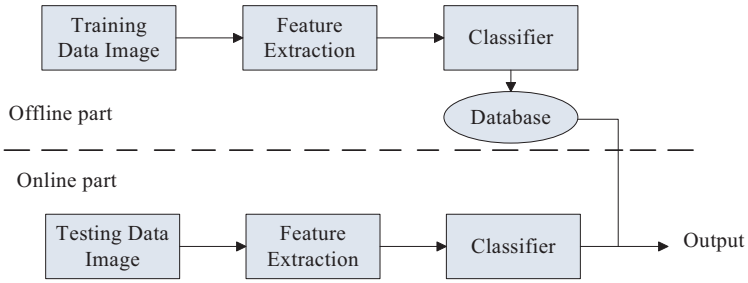


Fig. 111.1 The whole experiment process

loaded into support vector machine (SVM) for training and classification. SVM [4] is a trainable learning machine put forward by Vapnik et al. SVM is based on the Vapnik–Chervonenkis (VC) theory; it uses the structural risk minimization (SRM) principle and shows advantages in the small-sample, nonlinear, and high-dimension pattern recognition problem. The basic idea of SVM for pattern recognition is to define an optimal linear hyperplane to make the best generalization ability [5], and the research of the optimal linear hyperplane algorithm is summed up as solving a convex quadratic programming problem. Then, based on the Mercer kernel expansion theorem [6], the sample space is mapped to a high-dimensional even infinite-dimensional feature space through nonlinear mapping. In this feature space, we can utilize a linear learning to solve high-dimensional nonlinear classification in the sample space. The whole experiment process is shown in Fig. 111.1.

111.2 Face Recognition Based on PCA-SVM Algorithm

111.2.1 PCA Used for Feature Extraction

111.2.1.1 Traditional PCA

The purpose of PCA is to find a set of optimal-unit orthogonal vector basis by linear transformation. Then, the linear combination of the basis is used to reconstruct the original sample and to get the minimum mean square error (MSE) between the reconstructed and original samples. It can be proved that the projection vectors can be obtained by solving the eigenvalue problem in PCA [3].

$\mathbf{X}$ is defined as an $N \times M$ data matrix with column vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M$. The basis Φ of Karhunen-Leover transformation (KLT) is obtained by solving the eigenvalue problem $\Lambda = \Phi^T \Sigma \Phi$, where Σ is the data covariance matrix

$$\Sigma = \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{x}_i \boldsymbol{x}_i^T . \tag{111.1}$$

Fig. 111.2 The top five eigenfaces in ORL database

Therefore, the data must be projected onto Φ_j to get the first K main components by the PCA. A set of coordinate coefficients is obtained to serve as the basis for face recognition. The first K column vectors of the basis Φ are corresponding to the maximum K eigenvalues. Thus, the dimension of the face image decreases from N to K and meanwhile contains the largest energy of the signal. The coefficients of dimension projection are loaded into the classifier for the identification as the eigenvectors of the facial image. As shown in Fig. 111.2, the first five eigenfaces are obtained by PCA.

111.2.1.2 Fast PCA

For example, the Olivetti research laboratory (ORL) database consists of 40 people, and each has 10 different images. Each image is 112×92 pixel, namely $N = 10304$. The covariance of the data matrix is a $N \times N = 10304^2$ square matrix. It is rather difficult to calculate the eigenvector and eigenvalue directly. Then, the fast PCA algorithm is adopted in this chapter. We calculate the covariance matrix $\Gamma = \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{x}_i^T x_i$, which becomes $M \times M$ square matrix. Since the ORL database is a small-sample database, and the sample in training phase is less than the overall sample, the problem of calculation is solved perfectly.

111.2.2 SVM Used for Classification

Generally speaking, SVM is a kind of binary classification model. The basic model is defined as the largest margin linear classifier in the feature space. Namely, the learning strategy of SVM is to maximize the function margin. Eventually, the problem is transformed into a convex quadratic programming problem.

111.2.2.1 SVM Algorithm

We assume that the training sample set for $\{(\boldsymbol{x}_i, y_i), i = 1, 2, \ldots, p\}$ consists of two categories. It is marked as $y_i = 1$ if $\boldsymbol{x_i}$ belongs to the first category; it is marked as $y_i = -1$ if $\boldsymbol{x_i}$ belongs to the second category. The goal of learning is to construct a discriminant function that can classify the testing data as correctly as possible. The training sample set would be discussed separately as linear and nonlinear there after.

The Case of Linear

If there exists the optimal separating hyperplane

$$\boldsymbol{w} \bullet \boldsymbol{x} + b = 0. \tag{111.2}$$

According to the statistical learning theory, if the training sample is not mistakenly separated by the hyperplane, and the distance between the hyperplane and the sample data closest to the hyperplane is maximum, then this classification hyperplane is the optimal hyperplane.

Thus, the value of $\boldsymbol{w}$ should satisfy the maximization of $2/\|\boldsymbol{w}\|$. It can be transformed to a quadratic programming problem:

$$\begin{cases} \min\limits_{\mathbf{w},b} \dfrac{1}{2}\|\boldsymbol{w}\|^2 \\ s.t. y_i\left(\boldsymbol{w} \bullet \boldsymbol{x_i} + b\right) \geq 1, i = 1, 2, ..., p \end{cases}. \tag{111.3}$$

The optimal solution of Eq. (111.3) is solved by the Lagrange multiplier method:

$$L(\boldsymbol{w}, b, \alpha) = \frac{1}{2}\|\boldsymbol{w}\|^2 - \sum_{i=1}^{p} \alpha_i \left\{ y_i[(\boldsymbol{w} \bullet \boldsymbol{x}_i) + b] - 1 \right\}, \tag{111.4}$$

Where α_i are the Lagrange nonnegative coefficients. The Lagrange has to be minimized with respect to $\boldsymbol{w}$ and b, and maximized with respect to α_i. Equation (111.4) is solved by converting it to its dual problem [7], which greatly simplifies the solving process of Eq. 111.4. The dual problem is

$$\max_{\alpha} W(\alpha) = \max_{\alpha} \left\{ \min_{\boldsymbol{w},b} L(\boldsymbol{w}, b, \alpha) \right\}. \tag{111.5}$$

The solution of the dual problem is as follows:

$$\overline{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{p} \alpha_i - \frac{1}{2} \sum_{i=1}^{p} \sum_{l=1}^{p} \alpha_i \alpha_j y_i y_j \boldsymbol{x}_i \bullet \boldsymbol{x}_j, \tag{111.6}$$

$$\text{s.t.}\begin{cases} \sum_{i=1}^{p} \alpha_i y_i = 0 \\ \alpha_i \geq 0, i = 1,\ldots,p \end{cases}. \tag{111.7}$$

The situation discussed above is the training sample set of linear separable. And for the training sample set of linear inseparable, a nonnegative slack variable $\xi_i, i = 1,2,\ldots,p$ [8] is introduced. The optimization problem of classification hyperplane is

$$\begin{cases} \min\limits_{\mathbf{w},b} \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum\limits_{i=1}^{p} \xi_i \\ s.t.\ y_i\left(\boldsymbol{w}\bullet\boldsymbol{x_i} + b\right) \geq 1 - \xi, i = 1,2,...,p \end{cases}, \tag{111.8}$$

where C is a penalty coefficient and represents the penalty of misclassification.

The Case of Nonlinear

When the training sample set is nonlinear, the solution of SVM is to choose the kernel function method to solve the inseparable problem in the original space by mapping the data into the high-dimensional space [9].

The dual problem is

$$\max_{\alpha}\{L_D = \sum_{i=1}^{p} \alpha_i - \frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p} \alpha_i \alpha_j y_i y_j \Phi(\boldsymbol{x}_i)\bullet\Phi(\boldsymbol{x}_j)\}, \tag{111.9}$$

$$\text{s.t.}\begin{cases} \sum_{i=1}^{p} \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C, i = 1,\ldots,p \end{cases}, \tag{111.10}$$

Where $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \Phi(\boldsymbol{x}_i)\bullet\Phi(\boldsymbol{x}_j)$ is called the kernel function. There is no need of an explicit calculation of this nonlinear function in solving the optimal problem and calculating the classification function, thereby the curse of dimension of the feature space could be avoided [10]. The radial basis function $K(\boldsymbol{x}_i, \boldsymbol{x}) = \exp(-\|\boldsymbol{x}_i - \boldsymbol{x}\|^2 / \sigma^2)$ is applied in this chapter.

111.2.2.2 Parameter Optimization

The classification performance of SVM is dependent on the choice of parameter. It is not only time consuming but also not comprehensive to use manual testing in a traditional way. Literatures [11–13] reported several parameter optimization methods, including genetic algorithm (GA) [11], particle swarm algorithm (PSO) [12] and grid search algorithm (GS) [13]. These algorithms improve the recognition rate with the increase of the recognition time, which would pay more expenses in the practical application. Combined with the CV algorithm [14], the program, considering both recognition rate and recognition time, greatly reduces the recognition time.

The CV method is a commonly used method and is easy to implement. But, the calculation amount is large, especially for the large-sample problem. The principle of the CV method is to divide the training data set into K subsets. Each time the $K-1$ parts of data are used as the training data, and the remaining part of data are used as the testing data. The MSE obtained from k iterations is used for estimating the generalization error. Finally, a set of optimal parameters for SVM classification is selected.

111.3 Experimental Results

This experiment was conducted on the ORL database. The images in the ORL database are gray images, including changes in expression, tiny attitude, and less than 20% of the scale change. The feature was extracted, and the dimension was reduced by PCA. The data are classified by SVM a nonlinear kernel combined with CV. The experimental results on five training images are shown in Fig. 111.3. The results show that the dimension of the feature vector by PCA will affect the recognition rate. The recognition rate obviously increases as the dimension of eigenvectors increases from 9 to 80; however, the recognition rate increases slightly when the dimension of eigenvectors exceeds 80.

When the face image contains noisy data, for example, the salt and pepper noise, the accuracy rate increases to 90.5%; however, the error rate also increases to 1%. The explanation to them is the noisy data are also wrongly as face feature.

A comparison of PCA + SVM and PCA + SVM + VC is also conducted in this chapter. Take the first five, six, and seven images of each person as training data set for SVM training. The results are shown in Table 111.1. For five training images, the recognition rate of the proposed method is 89.5%, which is 5% higher than the PCA + SVMs. When the number of training images is seven, the recognition reached 96.67%. The results show the effect of the size of training sample set on classification. It can obtain a better eigenface space as the increasing of sample size, which will improve the classification in the subsequent testing stage. It indicates that the recognition rate of the integrated face recognition method proposed in this chapter is superior to the normal method.

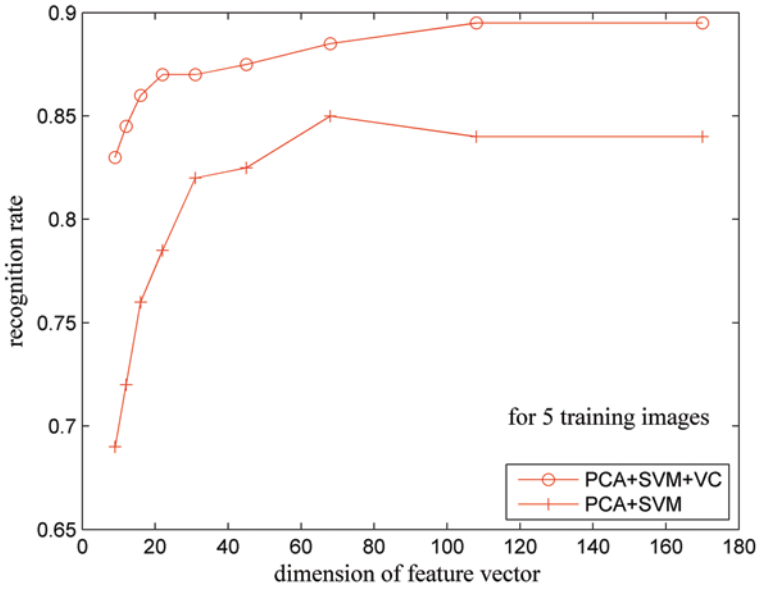


Fig. 111.3 The relation between the dimension and recognition rate

Table 111.1 Experimental result on ORL database

Algorithm	Training images	Testing images	Recognition rate (%)
PCA+SVM	5	5	84
PCA+SVM+VC	5	5	89.5
PCA+SVM	6	4	90
PCA+SVM+VC	6	4	96.25
PCA+SVM	7	3	93.33
PCA+SVM+VC	7	3	96.67

111.4 Conclusion

This chapter proposed a new method for face recognition using SVM nonlinear kernel coupled with CV. In the contrast experiments mentioned above, the significance of SVM as a face recognition classifier was verified. Experimental results also demonstrated the effectiveness of this face recognition system based on PCA + SVM coupled with VC. Thus, we can apply this face recognition system into the face autofocus of digital camera, access control system, credit card payment, and so on.

Acknowledgments This work is supported by the Research Projects of Zhengjiang Natural Science Foundations (R10900138) and the Zhejiang QianJiang Talents Project (2013R10071).

References

1. Li SZ, Jain AK. Handbook of face recognition. London: Springer; 2005. p. 19–691.
2. Noushath RS. Subspace methods for face recognition. Comp Sci Rev. 2010;4(1):1–17.
3. Jafri R, Arabnia HR. A survey of face recognition techniques. J Inf Process Syst. 2009;5(2):41–68.
4. Vapnik VN. An overview of statistical learning theory. IEEE Trans Neural Netw. 1999;10(5):988–99.
5. Maji S, Berg AC, Malik J. Classification using intersection kernel support vector machines is efficient. IEEE Conference on Computer Vision and Pattern Recognition, IEEE; 2008. p. 1–8.
6. Burges CJC. A tutorial on support vector machines for pattern recognition. Data Min Know Discov. 1998;2(2):121–67.
7. Guo GD, Li SZ, Chan K. Face recognition by support vector machines. IEEE Conference on Automatic Face and Gesture Recognition, IEEE; 2000. p. 196–201.
8. Yang Y, Ganesh A, Sastry SS, Yi M. Robust face recognition via Sparse Representation. IEEE Trans PAMI. 2009;31(2):210–27.
9. Scholkopf B, Burges CJC, Smola AJ. Advances in kernel methods: support vector learning. Cambridge: MIT Press; 1999. p. 169–84.
10. Gualtieri JA, Chettri S. Support vector machines for classification of hyperspectral data. Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International. IEEE; 2000. p. 813–5.
11. Wang SY, Tao L, Wang HB. Face recognition based on support vector machine and genetic algorithm. Comp Eng Appl. 2009;45(12):164–6.
12. Huang CL, Dun JF. A distributed PSO–SVM hybrid system with feature selection and parameter optimization. Appl Soft Comput. 2008;8(4):1381–91.
13. Li L, Zhang XL. Optimization of SVM with RBF kernel. Comp Eng Appl. 2006;29:190–2. (In Chinese).
14. Feng GH. Parameter optimizing for support vector machine classification. Comp Eng Appl. 2011;47(3):123–4. (In Chinese).

Chapter 112
The Management and Monitor System of Tunnel Construction Based on Internet of Things

Hong Cheng, Nan Wu and Jianwei Lian

Abstract A management monitor system of tunnel construction based on internet of things (IOT) is proposed. The system is to solve the problem of electrical connections instability caused by wired communications, the great difficulty of wireless network connections and the high cost, etc. And it is composed of a remote central control center, the main control unit, multiple data acquisition units and ZigBee terminal nodes. When the system is initialized successfully, the main control unit and the data acquisition units are connected to the remote central control center via wireless network by manual mode or automatic mode and it can also obtain the various sensors information of monitor system by sending control and query commands. The wireless network of WiFi, GPRS (General Packet Radio Service) and ZigBee are integrated together and ZigBee is arranged in a specific layout of tunnel in order to solve the above problem. The system that proposed in the paper has solved the above problems, reducing the probability of accidents and eliminating the security risks and it can also make a more rational scheduling and safety monitoring management of the on-site personal and the equipment.

Keywords Internet of things · WiFi · GPRS · ZigBee · Monitor system

112.1 Introduction

With the development of society, the pace of urbanization is faster and faster and the pressure of urban traffic is also growing. So, the underground tunnels are rapidly applied, because of its fast, convenience and no occupying ground area. Since the tunnels are constructed underground, the difficulty is hard to imagine and the normal schedule can easily affected by the engineer disasters. How to real-time and accurately obtain the personals and the constructions information is an effective measures to avoid the above problems [1].

H. Cheng (✉) · N. Wu · J. Lian
Science and Technology on Electromagnetic Compatibility Laboratory,
China Ship Development and Design Center, 430064 Wuhan, China
e-mail: hong027121@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_112

The management monitor system of tunnel construction based on IOT can solve the problems. The real-time networking and data transferring between the remote central control center and the main control unit are realized by the arrangement of WiFi and GPRS network in the tunnels. The system can switch to the GPRS network automatically [2–3], when the tunnel WiFi network fails or the signal cannot be covered, to achieve no blind messaging. The main control unit and the data acquisition unit are networked by ZigBee network and the data acquisition units are arranged throughout the tunnel by ZigBee network too, to ensure the real-time and accuracy of the system data.

112.2 ZigBee Network

ZigBee Network has the advantages of low cost, low power, low complexity, high network capacity and reliability. There are usually three ZigBee devices in ZigBee Network, including the coordinator node, which is the core equipment of ZigBee Network, the routing node, which is the repeater and the terminal node, which generally has the communication functions. The coordinator node and routing node must be the FFD device and each ZigBee Network can only have one coordinator node [4].

The data transfer direction of ZigBee Network is two-way and there exits three topologies: star network (Star), tree network (Tree) and mesh-type network (Mesh). The mesh network and AODV algorithm are adopted and its basic principle is to obtain routing lookup and record through dissemination routing packets which consist of routing request packet (RREQ) and routing reply packet (RREP) [5]. This method can increase the transmission depth of ZigBee Network, reduce the power consumption of ZigBee module and also improve the network transfer efficiency.

112.3 The System Design

112.3.1 The General Architecture Diagram

The management monitor system of tunnel construction based on IOT is composed of one main control unit that can be as the coordinator node of ZigBee network, multiple data acquisition units that the routing node and multiple ZigBee terminal nodes that the identification module wore on the body of the staff. The coordinator node is located on the entrance to the tunnel and the four routing nodes are distributed uniformly on the semicircle weeks, which the coordinator as the center and the maximum effective transmission distance of coordinator node device as the radius. Then, the routing nodes are located on the vertex of the equilateral triangle respectively, with the maximum distance as the side length of the triangle, shown in

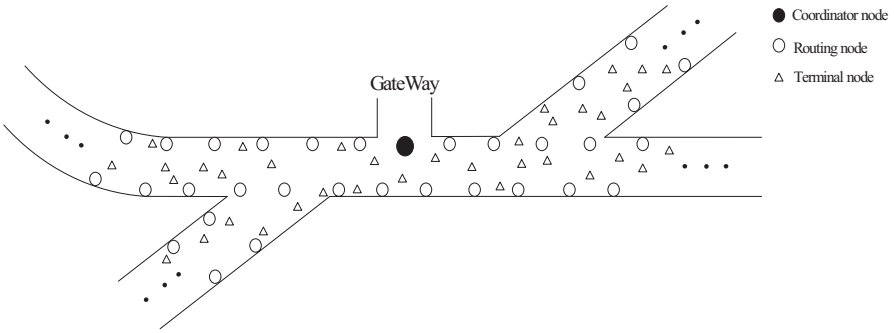


Fig. 112.1 The schematic layout diagram of ZigBee Network equipment in tunnels

Fig. 112.1. That arrangement can ensure that there are at least two adjacent routing nodes to transfer information, which can avoid the interruption of communication caused by the failure of ZigBee routing nodes and further strengthen the stability of ZigBee network. Besides, the simple structure make low design cost and can inspect the failed node rapidly.

112.3.2 The System Hardware Design

The system hardware mainly includes the remote central control center, the main control unit, the multiple data acquisition units and the multiple ZigBee terminal nodes, as shown in Fig. 112.2. As the human-computer interaction window of whole system, the main control unit includes wireless communication module, human-computer interaction modules, power modules, ZigBee network modules and MCU1 core modules of micro control unit. Wireless communication modules consist of GPRS and WiFi modules and are connected to MCU1 modules of micro control unit through serial port1 and port2, to realize the information transfer with remote central control unit. Human-computer interaction modules are connected to MCU1 modules by USB and can real-time display the status and fault information. ZigBee network modules are connected to MCU1 modules by serial port3 and achieve the information transmission with data acquisition units. The LED indicators, watchdog, clock, reset and EEPROM circuit make up of MCU1 modules of micro control unit.

As the information collection window of system, data acquisition unit includes power modules, MCU2 core modules and data acquisition modules. MCU2 core modules are composed of LED indicators, watch dog, clock and reset circuit. Data acquisition modules consist of temperature and humidity measurements, surface measurements and flammable gas measurements. It is connected to the terminal nodes via ZigBee wireless network. It receives the staffs' location information sent by the terminal nodes, and also is connected to data acquisition modules through ZigBee network.

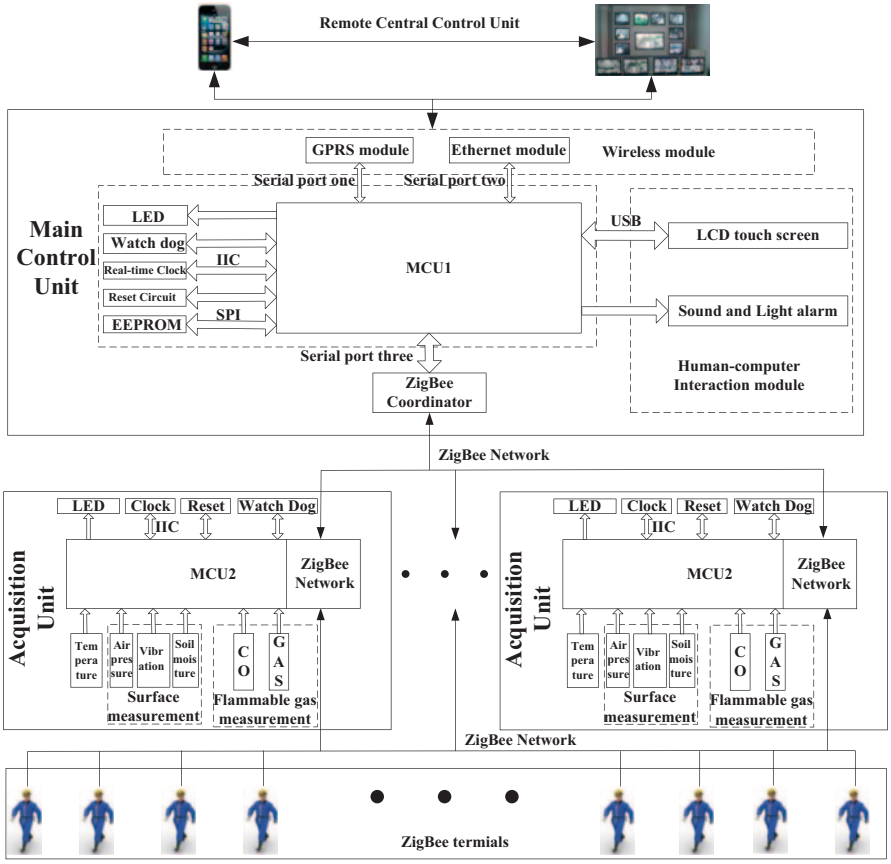


Fig. 112.2 The system hardware block diagram

ZigBee terminal nodes, which are the active electronic tags worn on the body of the staffs, are connected to the data acquisition unit and send the location information once in a while.

112.3.3 The System Software Design

When the system is initialized successfully, the main control unit and data acquisition unit are automatically or manually connected to the remote central control center via wireless network, as shown in Fig. 112.3. Then, the control center sends the control and query commands to obtain the information of sensors for tunnel construction management monitor system, the system status information and self-diagnostic information. The main control units analyze and process the received data, store the system fault code, the fault time, the process parameters and the op-

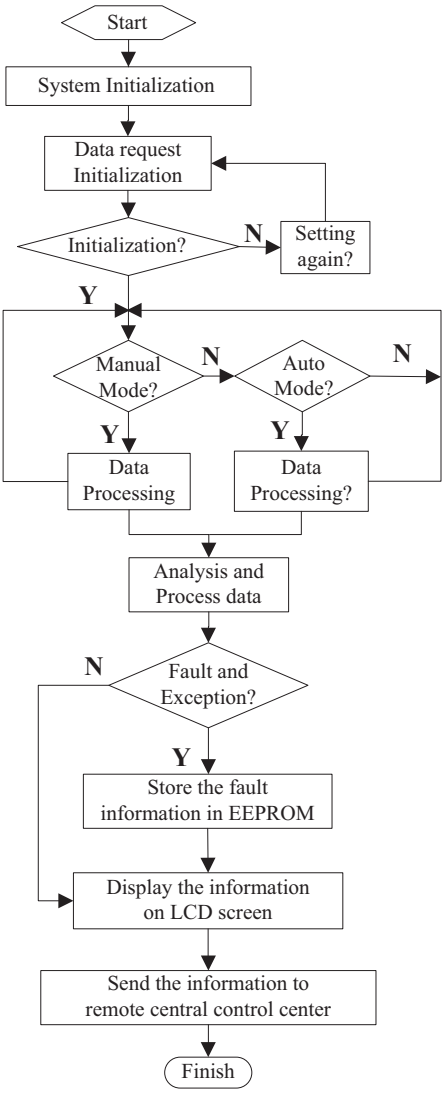


Fig. 112.3 The flow chart of system monitors method

erating status in EEPROM and also display the information on LCD touch screen. Finally, it sends the analyzed information to the remote central control unit through WiFi or GPRS network.

In manual mode, when the network is connected successfully, the main control unit sends connection report to the remote central control center, through WiFi or GPRS network selected by the remote central control center. Meanwhile, the control and query commands are sent to main control unit via the selected network. Once the commands are received, coordinator is started and the received commands are parsed. Then, the parsed commands are sent to data acquisition units and the

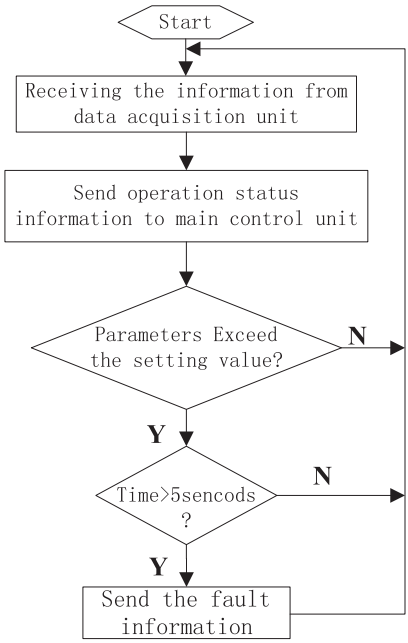


Fig. 112.4 The flow chart of data acquisition

routing notes are stared. The current temperature and humidity value, combustible gas concentrations, surface condition and staffs' positions are collected by the data acquisition units. Besides, the collected information is sent to the main control unit via ZigBee network if needed.

In automatic mode, the WiFi network is selected firstly by the remote central control center. If failed, the GPRS network is then connected and the connection report is sent by the main control unit. The other processes are the same with manual mode.

Shown in Fig. 112.4, when the detected temperature reaches 35 °C or the humidity reaches 70 %, the main control units send commands to sound and light alarm module and the buzzer alarms to alert the staffs carefully. Also, when the real-time temperature or humidity of the tunnel environment reaches 40 °C or 80 % respectively, the commands are sent to sound and light alarm module. Then, the light alarms and the fault locations are sent to alert the staffs to inspect the fault.

When the concentration of the detected combustible gas reaches 0.5 % or the CO reaches 200 ppm, the sound and light alarm module receives commands and the buzzers and fan equipment are started. When it reaches 1 % or 300 ppm, the light alarms and fault locations are sent to alert the staffs to inspect. When it reaches 2 % or 400 ppm, the sound and light alarms and the fault locations are sent. Meanwhile, the staffs stop work and evacuate and the professionals are requested to inspect the failures.

When the detected values of the acceleration sensors, the soil moisture sensors, or the air pressure sensors of the surface state information is more than 2 g, 60 % or

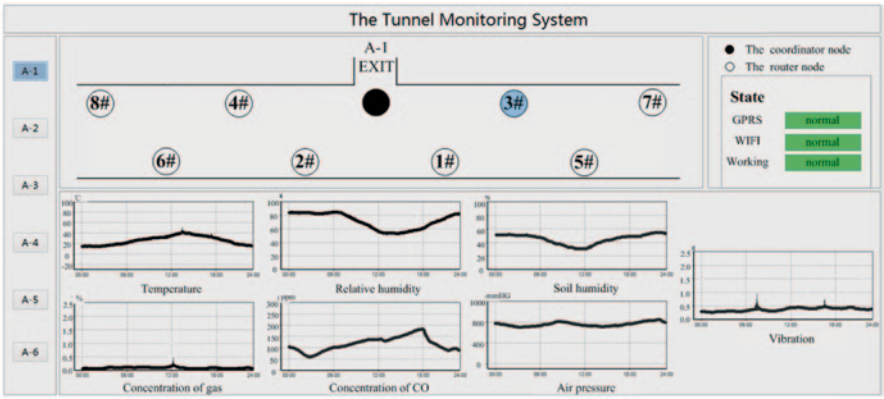


Fig. 112.5 The interface of PC software

770 mmHG respectively, the buzzers are started. If any two of them reach the value, the light alarms and the fault locations are sent to stop work and inspect the fault. If all of them reach the value, the sound and light alarm and all of the staffs evacuate and the professionals are requested to inspect the failures.

112.3.4 The System PC Software

The functions of PC software are to display the operation status of each main control unit and the obtained data information of the sensors. Besides, the detected data can be readily stored in text format, to long-term track, report, query and analysis facilitate. The interface of PC software is shown in Fig. 112.5.

112.4 Conclusion

The paper has briefly introduced the basic principle of management monitor system of tunnel construction based on IOT and the design process of hardware and software system. The PC software has also been designed to monitor the system. The designed system can greatly improve the reliability and real-time of the system. Above all, the system is cost-effective, easy maintenance and expansion node, which can increase the automation of tunnel construction management.

References

1. Bandyopadhyay D, Sen J. Internet of things: applications and challenges in technology and standardization. Wirel Pers Commun. 2011;58(1):49–69.
2. Roy F, Alex K, Yevgeny K. On power and throughput tradeoffs of wifi and bluetooth in smartphones. IEEE Trans Mob Computing. 2013;12(7):1363–76.
3. Lu S, Wang Y, Li Z, He D. Development of a GPRS-based environment monitoring system for tobacco production. Sens Lett. 2011;9(3):1–7.
4. Huang Y, Pang A, Hsiu P, Zhuang W, Liu P. Distributed throughput optimization for ZigBee cluster-tree networks. IEEE Trans Parallel Distrib Syst. 2012;23(3):513–20.
5. Cuomo F, Abbagnale A, Cipollone E. Cross-layer network formation for energy-efficient IEEE 802.15.4/ZigBee wireless sensor networks [J]. Ad Hoc Netw. 203;11(2):672–86.

Chapter 113
Research and Design of Lossy Compression Algorithm in Embedded Real-Time Database

Xinli Li, Hongkai Qiu and Yaochun Zhu

Abstract As the embedded real-time database (ERTDB) has high requirements on the compression algorithm in terms of compression time, compression ratio and compression precision, in this chapter, the dead zone algorithm, swing door tending algorithm (SDT), and the improved SDT algorithms are analyzed in depth, and a lossy compression algorithm is designed which can be applied to ERTDB. This algorithm is based on SDT with two improved methods, the elimination of wild value and the optimization of approximate line, combined in this algorithm. The test result shows that it can achieve good compression effect in ERTDB.

Keywords Embedded real-time database · Swing door tending · Compression time · Optimizing · Elimination

113.1 Introduction

In the data compression of embedded real-time database (ERTDB), there are three key problems. Firstly, high compression ratio is important because the hardware resources are limited and a huge amount of data has to be stored; secondly, high-speed performance of compression and decompression is required in the compression algorithm because the storage and the query of data is real time in ERTDB; thirdly, high compression precision is required in order to guarantee that the decompressed data can describe the trend of process accurately. Generally speaking, it is contradictory to meet the requirement of compression ratio, compression time, and compression precision simultaneously; therefore, it is necessary to find the proper balance among those requirements in order to achieve better performance for the application of data compression technology in ERTDB.

X. Li (✉) · H. Qiu
School of Control and Computer Engineering, North China Electric Power University, 102206 Beijing, China
e-mail: lixinli@ncepu.edu.cn

Y. Zhu
Longyuan (Beijing) Wind Power Engineering Technology Co., Ltd, 102206 Beijing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_113

The most common lossy compression algorithms in process data compression are dead zone algorithm (DZ) [1]and swing door tending algorithm (SDT) [1, 2]. There are also many improved methods for SDT, such as adding the forced storage recording limit [3–5], optimizing the approximation line [3, 4], adaptively adjusting the compression deviation[5–10], storing the point that has a great influence on compression precision [9], eliminating the wild value and dynamically adjusting the fitting function [10]. Although great effort has been devoted to the research of those algorithms, there is no research on whether they are suitable for ERTDB.

In this chapter, those lossy compression algorithms are analyzed in depth. The eliminate wild value and optimize approximate line SDT algorithm (EWOL-SDT) is designed and presented in order to meet the requirement of ERTDB on compression time, compression ratio, and compression precision of compression algorithm. EWOL-SDT algorithm may improve the approximate line and make up the drawback of SDT in wild value detection. The experimental and practical application results show the feasibility.

113.2 Lossy Compression Algorithm Analysis

113.2.1 Dead Zone Algorithm

The principle of DZ is as follows. A threshold is set in advance. If the deviation between the current data and the last stored data is larger than the threshold, the current data will be stored [1]. The compression precision of DZ is low so that the data cannot describe the tendency of process accurately after decompression. It has a bad compression result on the linear drift data.

113.2.2 SDT Algorithm

The essence of SDT algorithm is to use a line that is determined by two adjacent stored points instead of other data between the two stored points [1, 2]. In Fig. 113.1, the last stored point is A, the distance between point A and point A_1 is the same as the distance between point A and point A_2 in the vertical direction. It is called the compression deviation (ΔE) and can be set artificially. A parallelogram whose length of side is double of ΔE and through point A and the current point is designed. If the parallelogram contains all points between point A and the current point, the compression will be continued; otherwise, the previous point will be stored. As shown in Fig. 113.1, point D will be stored and used with line AD instead of A, B, C, and D.

Obviously, SDT algorithm has short calculation periods and is easy to implement. It also has high compression efficiency and compression deviation because it can be controlled; however, it is unable to detect and handle the wild value.

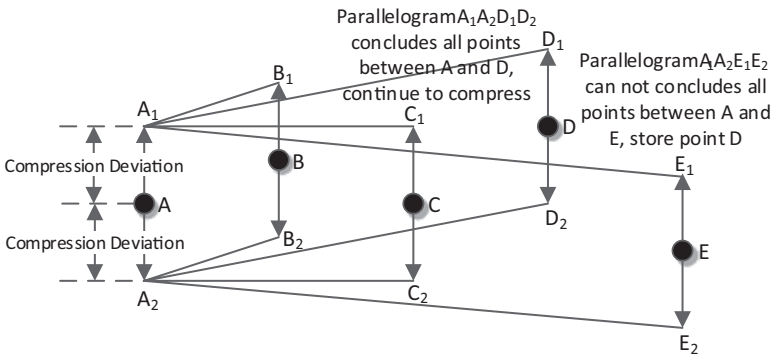


Fig. 113.1 Schematic diagram of SDT algorithm

113.2.3 Improved SDT Algorithm

113.2.3.1 Forced Storage Recording Limit (FSRL)

No data can be stored for a long time if a measured point changes very slowly when the SDT algorithm is used to compress data. This will affect the accurate description of process trend [3–5]. In order to avoid such a situation, the forced storage recording limit (FSRL) can be set in advance. Once the length of current round of compression interval meets the FSRL, the current point will be saved.

113.2.3.2 Optimize Approximation Line SDT (OL-SDT)

SDT uses the line, which connects two stored points, to represent all data between them. In this way, the points which are compressed do not work [3, 4]. In order to enhance the importance of the compressed points, an approximation line is designed at the end of every round of compression. It has an average slope of the maximum slope and the minimum slope. The last stored point is used as the starting point.

113.2.3.3 Adaptively Adjusted Compression Deviation SDT (AAD-SDT)

The ΔE of SDT is preset as a fixed value. Large ΔE will lead to a high compression error; otherwise, the compression ratio will be very low. A better compression effect will be achieved if the ΔE can be adjusted dynamically according to some indicators. There are three types of indicators. One is the result of the comparison between compression interval length (CIL) and FSRL. This algorithm is denoted by AAD-SDT1 [5, 6]. The second is the result of the comparison between the actual compression average error δ and the expected compression average error δ_E. This algorithm is denoted by AAD-SDT2 [7, 8]. The third is the result of the comparison

between the standard deviation of the original data in current compression interval and in previous compression interval. This algorithm is denoted by AAD-SDT3 [9, 10].

113.2.3.4 Store Great Influence Point SDT (Max-SDT)

During the compression, some points greatly differ from other points. If these points are abandoned by SDT during compression, the compression precision will reduce dramatically; therefore, if we store these points, the compression precision will increase [9]. We set an expected record limit (RL) in advance to find the point that has a maximum compression error after a round of compression. If the maximum compression error exceeds the RL, the found point will be stored.

113.2.3.5 Eliminate Wild Value SDT (EW-SDT)

There are wild values (noise) in the actual data. They cannot be detected by SDT. Therefore, the wild value will be processed as normal data, which will not only increase the data compression ratio but also affect the accurate description of process trend [10]. In order to detect the wild value, a noise limit (NL) should be set. If the deviation of a data with both its adjacent points is larger than the NL, the data is a wild value and will be ignored.

113.2.3.6 Dynamically Adjust Fitting Function SDT (AF-SDT)

A straight line is used as the fitted curve in SDT. But in most cases, the data does not change linearly. If the changing trend of data is similar to a parabola, it will be better to fit the data with a shape of a parabola. If the changing trend of data is more approximate to the linear change, it will be better to use straight line as the fitting curve. In order to achieve a better compression effect, we can dynamically adjust the fitting function according to the changing trend of the actual data during compression [10].

113.2.4 Algorithm Performance Analysis

In order to test the performance of these algorithms above, the algorithms are implemented by C language. FSRL plays an important role in describing the data trends accurately, thus this paper will use it as a part of SDT. We use 15,222 actual current data obtained from small current earthed system for test, ΔE is set to 5, and FSRL is set to 20. As for AAD-SDT, the range of ΔE is from 3 to 7, δ_E of AAD-SDT2 is set to 2.5; the RL of Max-SDT is set to 3, and for EW-SDT, the NL is 50. Test results are shown in Table 113.1.

Table 113.1 Test result of algorithm

Compression algorithm	Compression ratio (CR)	Compression average error (CE)	Compression time (ms)
DZ	1.97	15.93	38
SDT	4.77	2.28	26
OL-SDT	4.73	2.02	28
AAD-SDT1	6.56	3.03	26
AAD-SDT2	7.21	2.51	91
AAD-SDT3	5.36	2.46	59
Max-SDT	4.54	1.34	60
EW-SDT	4.92	2.01	26
AF-SDT	5.55	2.09	84

As it can be seen, dead zone algorithm has a bad performance in every indicator. While compared with SDT, the CR of OL-SDT remains basically unchanged and the CE reduces by 9.1%; AAD-SDT1 improves the CR by 37.5% while the CE increases 32.9%; AAD-SDT2 improves the CR by 51.2% and the CE increases 9.2%; AAD-SDT3 improves the CR by 12.4%, at the same time, the CE increases 7.9%; the CR of Max-SDT is decreased by 4.8% with the CE rate by 41.2%; EW-SDT improves the CR by 3.1% and the CE reduces 9.4%; the CR of AF-SDT increases 16.4% while the CE reduces 5.9%. In respect of the compression time, OL-SDT, AE-SDT1 and EW-SDT use the same time as SDT; the compression time of other algorithms is more than twice as that of SDT.

113.3 Design of Lossy Compression Algorithm in ERTDB

Lossy compression algorithm for ERTDB needs to achieve a good performance on time, space, and compression precision, which need to be taken into consideration in the design.

Based on the above-mentioned analysis, the dead zone algorithm cannot achieve a good result in the compression of process data with linear drift characteristics; thus it cannot meet the requirements of ERTDB. SDT algorithm takes a low time-cost and gives a good performance in CR and CE so that it can be used in the embedded real-time database. Compared with SDT algorithm, the CR or the compression precision of AAD-SDT2, AAD-SDT3, Max-SDT, and AF-SDT has been improved; however, the compression time of them is too long to meet the requirements of ERTDB; OL-SDT algorithm can reduce CE without the CR upping. Although the compression time of OL-SDT has increased, it can still meet the requirements of ERTDB. The CR of AAD-SDT1 has been significantly improved, but its CE increases a lot; as a result, the trend of process cannot be described accurately, and it cannot be used in ERTDB. EW-SDT algorithm can improve CR and CE while the

compression time does not increase; what is more important, it can make up the drawback of SDT that cannot detect and handle the wild value, so it can be used in ERTDB.

As noted from the analysis above, the lossy compression algorithm used in ERTDB can adopt SDT with optimize approximation line and eliminate wild value, named as EWOL-SDT. EWOL-SDT improves the approximation line by an average slope of the maximum slope and the minimum slope at the end of every round of compression and eliminates the wild value through a noise limit.

D_{Cur} is the value of the current point, D_{Last} is the value of the last point, D_{BeLast} is the value of the before last point, D_{Stored} is the value of the stored point. We define:

$$S_{\max} = \left\{ (D_{\text{Last}} - D_{\text{Stored}} - \Delta E) / T_{\text{Dif}}, S_{\max} \right\} \tag{113.1}$$

$$S_{\min} = \left\{ (D_{\text{Last}} - D_{\text{Stored}} + \Delta E) / T_{\text{Dif}}, S_{\min} \right\} \tag{113.2}$$

$$S = (D_{\text{Last}} - D_{\text{Stored}}) / T_{\text{dif}} \tag{113.3}$$

Where, $S_{\max}$ is the maximum up slope of the last point and the stored point, $S_{\min}$ is the minimum below slope of the two points, S is medium slope of the two points, T_{Dif} is the time deviation between two points. The average slope S_{Avg} of the maximum up slope and the minimum below slope is shown as follows:

$$S_{\text{Avg}} = (LastS_{\max} + LastS_{\min}) / 2 \tag{113.4}$$

where $LastS_{\max}$ is the last $S_{\max}$ of the current compression, $LastS_{\min}$ is the last $S_{\min}$ of the current compression. The steps of the algorithm are as follows:

Initialization: initialize the parameters such as $\Delta E, NL,$ and FSRL;

Step1: If the point is the first point of the current round of compression, store it and denote it by D_{BeLast} and D_{Stored}, finish the compression; otherwise, go to Step 2;

Step 2: If the point is the second point of the current round of compression, denote it by D_{Last}, finish the compression; otherwise, go to Step 3;

Step 3: Denote the point by D_{Cur} if D_{Last} is a wild value judging by $D_{\text{BeLast}}, D_{\text{Cur}}$, and NL, go to Step 8; otherwise, go to Step 4;

Step 4: If the length of current compression interval is longer than FSRL, go to Step 6; otherwise, go to Step 5;

Step 5: Calculate $S_{\max}$ by Eq. (113.1), $S_{\min}$ by Eq. (113.2), S by Eq. (113.3). If $S_{\min} < S < S_{\max}$, go to Step 7; otherwise, go to Step 6;

Step 6: calculate S_{Avg} by Eq. (113.4), calculate the point to store by S_{Avg} and the time deviation of the before last point and the stored point, store the point and denote it by D_{Stored}, go to Step 7;

Table 113.2 Test result of EWOL-SDT

Compression algorithm	Compression ratio (CR)	Compression average error (CE)	Compression time (ms)
EWOL-SDT	4.89	1.88	28

Step 7: $LastS_{\max} = S_{\max}, LastS_{\min} = S_{\min}$, go to Step 8;
Step 8: $D_{\text{BeLast}} = D_{\text{Last}}, D_{\text{Last}} = D_{\text{Cur}}$, finish the compression.

113.4 Performance Test

The same original data are used to test EWOL-SDT, which is designed in this chapter, where, ΔE is set to 5, FSRL is set to 20, NL is set to 50. The test results are shown in Table 113.2.

As the result indicates, while compared with SDT, EWOL-SDT improves the CR by 2.5 %, reduces CE by 15.3 %. The compression time scales up by only 2 ms. It also makes up the drawback of SDT that cannot handle the wild value. Although the compression time of EWOL-SDT has increased, it is still more suitable for ERTDB than SDT considering its better performance than SDT on CR and CE. Compared with OL-SDT and EW-SDT, EWOL-SDT improves CR and CE while the compression time does not increase. As a result, it can achieve better compression effect in an ERTDB than OL-SDT and EW-SDT. Compared with AAD-SDT1, the CR of EWOL-SDT is reduced by 25 % but the CE is increased by 40.0 %. Because of high requirement of ERTDB on CE and accurate description of process trend, EWOL-SDT will achieve better compression precision than AAD-SDT1 in an ERTDB. EWOL-SDT has a great advantage of AAD-SDT2, AAD-SDT3, Max-SDT, and AF-SDT in compression time, which shows that it is more suitable for ERTDB than the others.

113.5 Conclusion

Based on the characteristics of ERTDB, which have high requirements on compression time, compression ratio, and compression precision, ENOL-SDT algorithm has been designed to meet the requirements of ERTDB in this paper. On the premise of guaranteeing the compression time, EWOL-SDT algorithm can improve compression ratio and compression precision; EWOL-SDT algorithm was applied to ERTDB platform researched by Weihan Technology Co., Ltd and was shown to achieve a good effect.

Acknowledgements This work is supported by the Fundamental Research Funds for the Central Universities (13MS21, 13MS24).

References

1. Wang J. Research and improvement of loss linear compression algorithms in the real-time database. Microcomput Appl. 2011;32(7):13–18. (In Chinese).
2. Zhen C, Ren B. Design and realization of data compression in real-time database. Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on, IEEE, 2009; 1–4.
3. Xing R, Qi Q, Zheng T. Improved SDT algorithm. Comput Eng Des. 2013;34(2):515–8. (In Chinese).
4. Silva IMD. Performance evaluation of a compression algorithm for wireless sensor networks in monitoring applications. Emerging technologies and factory automation, 2008. ETFA 2008. IEEE International Conference on, IEEE, 2008; 672–8.
5. Zhang W, Chen X. Research and application of improved process data compression algorithm-SDT. Ind Control Comput. 2009;22(8):1–3. (In Chinese).
6. Hale JC, Sellars HL. Historical data recording for process computers. Chem Eng Prog. 1981;77(11):38–43.
7. Qu Y, Wang W. Automatic parameter control SDT algorithm for process data compression. Comput Eng. 2010;36(22):40–2. (In Chinese).
8. Bristol EH. Swinging door trending: adaptive trend recording. Proceedings of the ISA '90 International Conference, ISA Services Inc, 1990; 749–53.
9. Yu S, Wang X, Zhao L. An improved swinging door trending (SDT) compression algorithm based on the tolerance dynamic adjustment. J Beijing Univ Chem Technol (Nat Sci). 2013;40(3):109–13. (In Chinese).
10. Ning H. A new process data compression algorithm based on SDT algorithm. Comput Technol Dev. 2010;20(1):25–8. (In Chinese).

Chapter 114
Multiple Channel Assignment and Recycling Algorithm for MR-MC Network Architecture

Ping Ren

Abstract In this paper, a channel assignment and recycling algorithm have been proposed, which assign wireless channels to the neighboring subnets to reduce the interference, meanwhile the network capacity has been improved by channel multiplexing between the far-away subnets and recycles channel when the subnet died out. The OPNET simulation results show that, it has a good performance in the aspects of network delay and throughput.

Keywords MR-MC · Wireless channel · Algorithm

114.1 Introduction

Wireless sensor network adopts the way of self-organization and multi-hop for communication, when all sensor nodes work at the same wireless channel, on one hand, it is hard to solve the problem of exposing and hidden terminal in essence; on the other hand, with the increase of sensor nodes, most of sensor nodes withdraw randomly as channel collision, and this will not only reduce the utilization ratio of channel, but also will decrease the network real-time performance and throughput [1].

With the development of wireless communication and system integration, the cost of wireless transceivers has been reduced greatly, so it is possible to equip multiple radios on sensor nodes, especially on backbone nodes (such as routers). So the MR-MC (Multi-radio Multi-channel) sensor network is the development direction of the future wireless sensor network.

The author has proposed an MR-MC manufacturing sensor network architecture based on the nested cellular topology structure in the previous research, the MR-MC network can be transformed into multiple SR-SC (Single radio Single channel) by subnetting, and this not only can perform the advantages in network capacity and channel access delay brought by multiple radios and multiple channels to wireless

P. Ren (✉)
Science and Technology on Electromagnetic Compatibility Laboratory,
China Ship Development and Design Center, 430064 Wuhan, China
e-mail: rp-2003@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_114

sensor network, but also keep pace with the current sensor network utmost, so the complexity brought by modifying MAC protocol has been reduced, and the existing research results can be made full use of [2].

114.2 Research Background

In order to solve the problems of competition and collision, researchers have introduced many multi-channel algorithms which mainly focus on modifying the current MAC protocols, but it is too complex and the delay brought by channel switching will affect the real-time performance of the network.

Ramachandran et al. proposed the BFS-CA (Breadth First Search-Channel Assignment) algorithm [3]. In this algorithm, the coordinator creates multi-radio conflict graph (MCG) based on the neighboring node information, single-hop delay and channel interference value, and calculates the distance between every apex and the coordinator in MCG. When accomplishing calculating, starting from coordinator, after visiting the coordinator, BFS-CA visit every unvisited neighboring nodes in graph by turns to ensure the routers nearest to the coordinator having high channel assignment priority, and then starting from the neighboring nodes and visiting their neighboring nodes in turn until the neighboring nodes of all the visited neighboring nodes in graph have been visited, and this make the early visited neighboring nodes prior to the late visited neighboring nodes. In order to guarantee the network packets transceiving during the channel switching, every node needs to reserve a special radio working at Radio Frequency Monitoring (RFMON) model and this has reduced the utilization ratio of radio.

KOB etc. proposed that the Distributed Channel Assignment (DCA) algorithm choose the wireless channel having minimum interference value according to the local topology information to realize the maximization of channel utilization ratio [4].

It can be seen from the above analysis that every channel assignment algorithm has its own advantages and disadvantages, while the sensor network having a high-relevance to application background, moreover the manufacturing environment has special performance requirements. Consequently, the multi-radio multi-channel assignment algorithm applicable to the manufacturing sensor network has the following characteristics:

1. The manufacturing sensor network is a control network and has high requirement on real-time performance, while the traditional multi-radio multi-channel assignment algorithm has introduced extra network delay as wireless channel switching. Consequently, during the life cycle of sensor node subnet, every subnet has to adopt preassignment and fix assignment strategy;
2. Because of the limitation of wireless channel resources, even equip multiple radios on router, the shortage of channel number will restrict greatly the scale and the capacity of the sensor network. Because the single-hop communication distance of radio is limited, the neighboring subnets adopt different communication

frequency and far away subnets reuse wireless channel will improve the network capacity greatly, and decrease the interference between them;
3. Because the time of subnets formed is not same, when the network runs stably, and then the subnet will produce, break up and die out as the mobility or state change of sensor nodes, so the channel assignment is a persistent and order assignment process, and this requires that the channel assignment algorithm has extendibility and can recycle the idle channel resource as subnet died out;
4. This network is a mixed network, routers equip multiple radios, sensor nodes choose one of the neighboring routers as its central node according to the received signal strength from them and the number of sensor nodes in subnet, and then use single radio to communicate within subnet, so channel assignment algorithm mainly focus on the multiple radios and multiple channels on routers actually, and has no direct relationship with the sensor nodes themselves;
5. The nested cellular topology is based on MR-MC adopt layer network design plan. Routers use special radio to communicate with the coordinator, and the top network has been assigned a fixed wireless channel, so it has special advantages in the aspect of network connectivity;
6. The manufacturing sensor network has been divided into multiple sensor node subnets, and controlling the number of sensor nodes in subnet and changing the method of wireless channel competition on sensor node can effectively reduce the end-to-end network delay;
7. In manufacturing network, their locations are fixed relatively when the equipment have been disposed, the locations of sensor nodes equipped on the equipment are fixed relatively, their mobile range is restricted and can be predicted even move, thereby the topology structure of the entire network has no major change during a period of time.

The current algorithms cannot well satisfy the requirements of manufacturing sensor network, so it is great significant to propose a channel assignment and to recycle algorithm according to the proposed network architecture.

114.3 The Proposed Algorithm

The topology of the nested cellular network proposed in the previous research by the author is shown in Fig. 114.1 [2], and this paper has proposed a new channel assignment algorithm.

It can be seen from Fig. 114.1 that sensor nodes adopt single radio to communicate within a subnet formed with the router according to the strength of wireless signal received from the router. The neighboring subnets use different channels to communicate within them, and the far away subnets reuse wireless channels to improve their utilization ratio and to enhance the network capability. While the communication between subnets and coordinator is transmitted by routers. The routers use another radio and channel to communicate with each other and coordinator, if the network scale is large enough, that can be accomplished by different radios. So

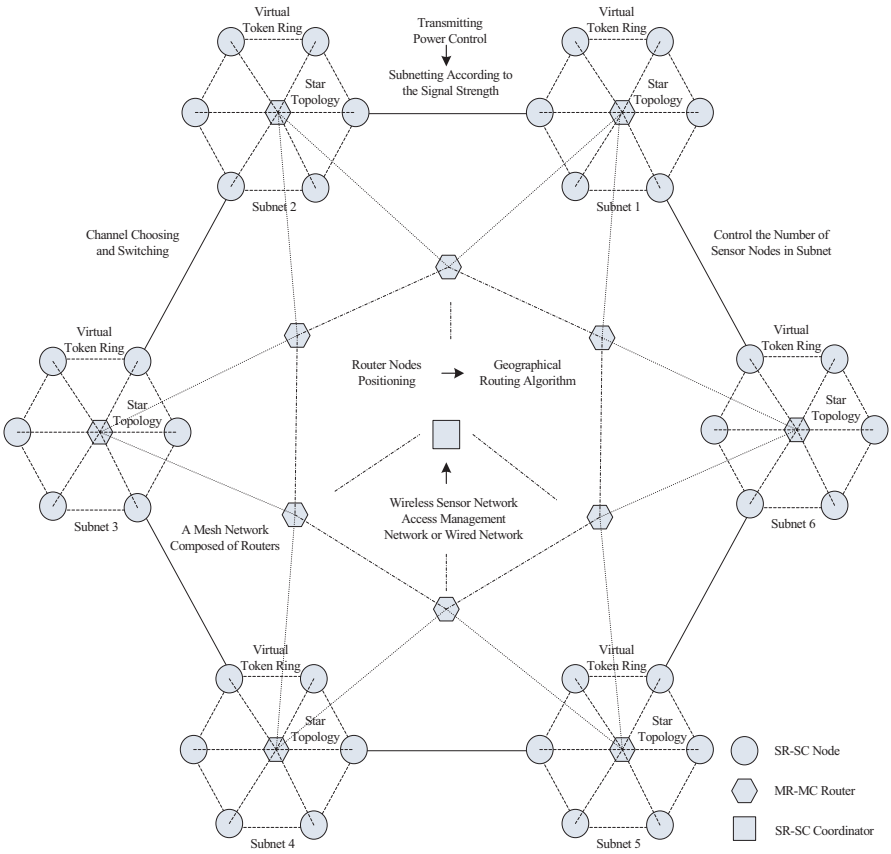


Fig. 114.1 MR-MC manufacturing sensor network architecture

the proposed multiple channel assignment and recycling algorithm mainly focuses on the multiple channels owned by the multiple radios on routers.

On account of the proposed network architecture by the authors in the previous research, channel assignment mainly assign channel for the multiple radios on router and make the neighboring subnets using different wireless channels to communicate within them, while the far away subnet can reuse the wireless channel to improve the network capacity, layer-networking can guarantee the connectivity of network, moreover controlling the number of sensor nodes in subnet and changing the competition method of wireless channel on sensor node can effectively reduce the end-to-end delay, and this can realize the goal of load equalization, decreasing transmission interference and algorithm complexity and having high extendibility. When the number of sensor nodes in subnet is less or more than minimum or maximum as the mobility or fault of sensor nodes, the subnet will die out or break up, at this time it is needed to recycle wireless channel from the nonexistent subnet or assign wireless channel for the new subnet, and this can effectively avoid the

"Ripple Effect" brought by channel assignment and the network delay brought by the frequent channel switching.

The ideal of the proposed algorithm is: the sensor network has been divided into different layers according to the number of radios on routers, and the number of layers is less than that of radios, which guarantee that different radio works at different network layer. When the network scale is small, a part of radios can stay in the idle state to save energy and improve efficiency. At the stage of network initialization, every sensor node starts and switches to its default channel. The coordinator acts as Channel Assignment Server (CAS), and then assigns wireless channel on account of the layer of wireless channel applying for and order according to the parameters passed from routers including the number of radios, neighboring routers and channel frequency, the neighboring subnets will be assigned to different channels, and to ensure that the interference between them is minimum as far as possible, while the far away subnets can be assigned to the same channel. During the life cycle of sensor node subnets, they will always use the assigned wireless channel to communicate within them to reduce the network delay brought by channel switching; when the subnet dies out, the coordinator recycles the wireless channel the subnet used to prepared for next assignment. As adopting layer model, the connectivity of sensor network can be fully guaranteed, while the minimization of interference between radios and the maximization of network capacity is the final goal of this channel assignment algorithm.

This proposed algorithm makes full use of the characteristics of the sensor network architecture under manufacturing environment, according to the channel interference and frequency reuse theory, not only ensure that the neighboring channels will not produce interference between them, but also can make full use of channel resource and improve the network capability greatly. The introduction of warning mechanism to channel assignment can effectively avoid the failure of channel assignment as there is no channel left for assignment, and improve the fault tolerance and robustness. Using the interference radius to judge whether existing interference between channels of the neighboring subnets are more accurate and persuasive than one-hop model and two-hop model, and this has shielded the difference between single hop distances as the unbalanced distribution of sensor nodes, otherwise it may affect the quality of wireless communication as the neighboring subnets will produce interference.

114.4 Simulation Experiments and Results Analysis

114.4.1 Simulation Model

OPNET is used to make a comparison between the proposed algorithm and two algorithms: DCA and BFS-CA [5]. The simulation model is shown in Fig. 114.2.

The simulation parameters are shown in Table 114.1.

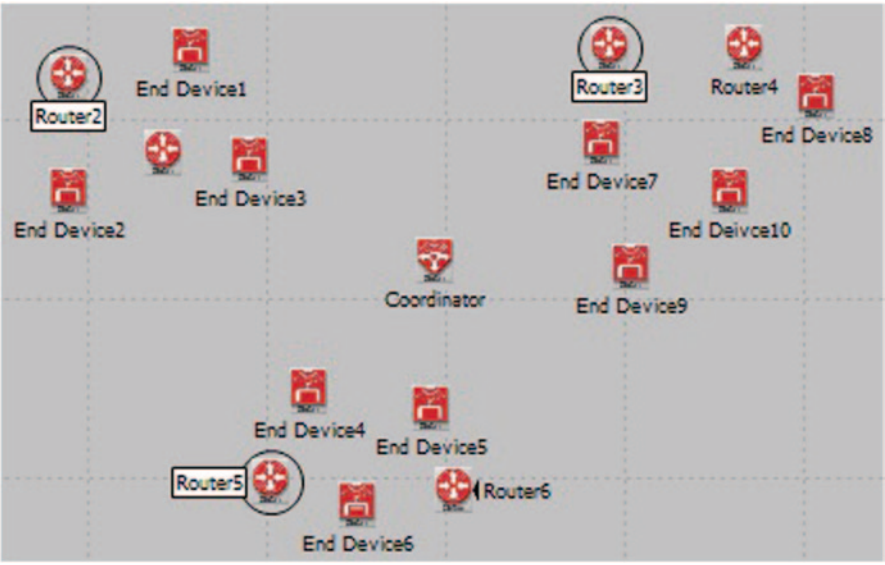


Fig. 114.2 Simulation model

Table 114.1 Simulation parameters

Simulation time	1200s	Number of coordinators	1
Number of sensor nodes	10	Number of radios on router	2
Number of routers	5	Number of channels	16

114.4.2 Results and Analysis

The comparison of MAC delay under the three channel assignment algorithms is shown in Fig. 114.3.

It can be seen from Fig. 114.3 that, DCA algorithm has the maximum of MAC delay and the most fluctuation; The amount of delay and the amplitude of fluctuation under BFS-CA algorithm and the proposed algorithm are similar, and they have no obvious performance difference [6].

The comparison of MAC throughput under the three channel assignment algorithms is shown in Fig. 114.4.

It can be seen from the Fig. 114.4 that, DCA algorithm has the least network throughput which has sharp down pulse usually. The network throughput of BFS-CA algorithm takes the middle place and has a process of jump at the time of network initialization. The proposed channel assignment algorithm has the maximum throughput when the network runs stably and hardly, and has the phenomenon of fluctuation [6].

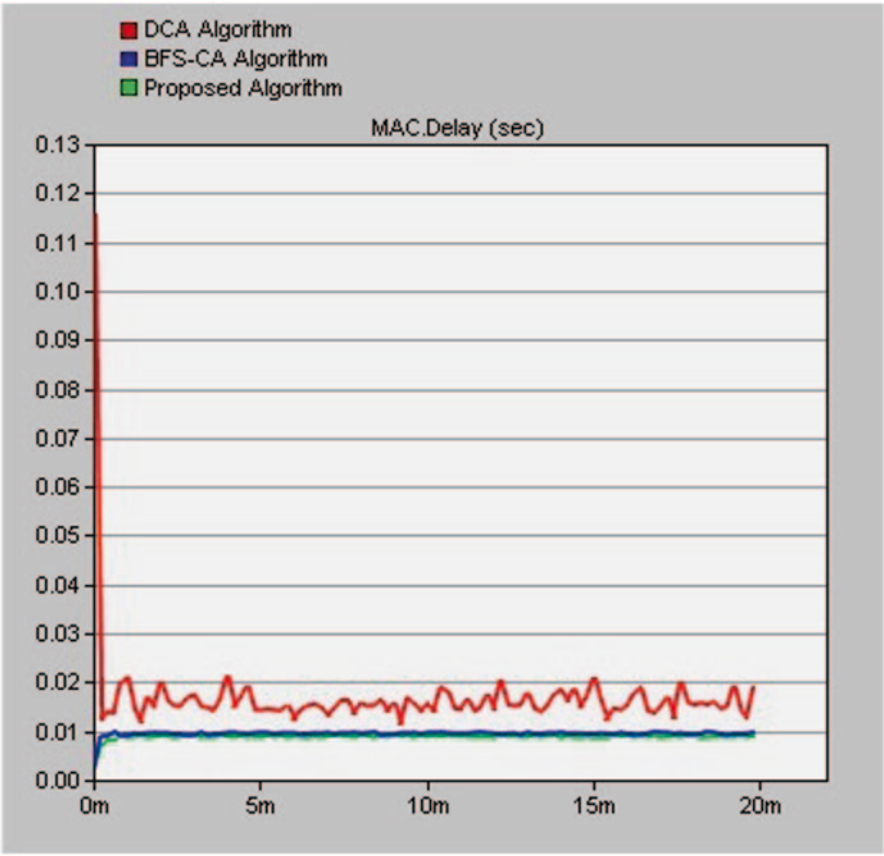


Fig. 114.3 The comparison of MAC delay

114.5 Conclusion

This paper firstly analyzes the characteristics and the requirements of the multiple channel assignment algorithm of MR-MC sensor network under manufacturing environment, and then proposes a centralized channel assignment algorithm based on requirements and topology structure. The OPNET simulation results show that, this algorithm is better than DCA and BFS-CA algorithm and has a good performance in the aspects of network delay and throughput.

Acknowledgements This paper is supported by the Fund of Key Laboratory of Science and Technology on Electromagnetic Compatibility.

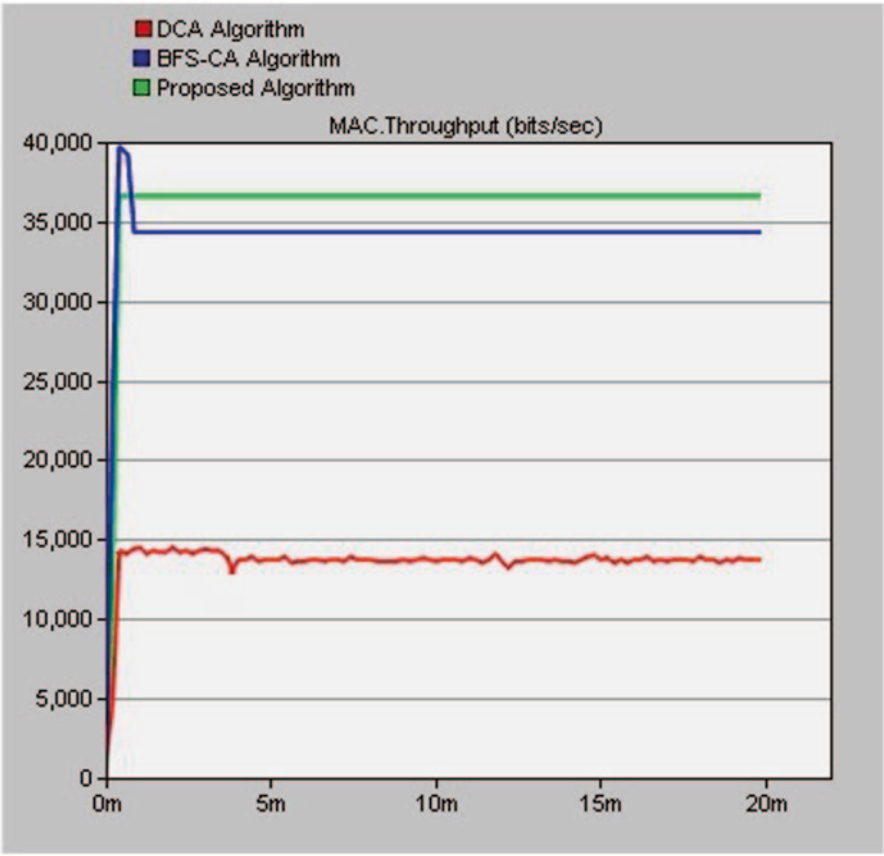


Fig. 114.4 The comparison of MAC throughput

References

1. Sun L. Wireless sensor networks. Beijing: Tsinghua University Press; 2005. p. 25–37 (In Chinese).
2. Ren P. Manufacturing sensor network based on multi-radio and multi-channel. Proceedings of 2014 Asia-Pacific conference on antennas and propagation, IEEE, in press; 2014. p. 1–8.
3. Ramachandran KN, Belding EM, Almeroth KC. Interference-aware channel assignment in multi-radio wireless mesh networks. Proceedings of 2006 25th IEEE International Conference on Computer Communications, IEEE; 2006. p. 1–12.
4. Ko B, Misra V, Padhye J. Distributed channel assignment in multi-radio 802.11 mesh networks. Proceeding of 2007 wireless communications and networking conference, IEEE; 2007. p. 1–11.
5. Min C. Opnet network simulation. Beijing: Tsinghua University Press; 2005. pp. 97–123.
6. Ren P. Multiple channel assignment and recovery algorithm for MR–MC network. Proceedings of 2013 IEEE Eighth international conference on networking, architecture and storage, IEEE; 2013. p. 232–6.

Chapter 115
A Social-Network-Based Intelligent Collaborative Decision Model with Larger Group Consensus Requirements for Collaborative B2B E-commerce

Xiang Liu

Abstract Collaborative decision technologies are crucial to support the growing complexity and diversity of the collaborative B2B e-commerce with larger group consensus requirements. The predefined group consensus is reached through several discussion rounds in a classic group decision-making by consensus models. However, the existing consensus models for the group decision with a low number of groups may not be the most appropriate approach to the collaborative decision with larger groups in collaborative B2B e-commerce context. In this paper, we present a social-network-based intelligent collaborative decision model with larger group consensus requirements for collaborative B2B e-commerce, which automates and supports the collaborative decision consensus-reaching processes. Finally, an example of the notebook on-line selling process is illustrated.

Keywords Collaborative decision · Group consensus · Collaborative B2B E-commerce · Social network

115.1 Introduction

Collaborative decision with larger group consensus requirements in collaborative B2B e-commerce context is a time-consuming process because all parties desire to maximize their own payoff while they may have common goals [2, 4, 5, 8, 9]. If some of the parties do not concede, it could take forever to reach consensus in a collaborative decision process. The approach to aggregate individual opinions into a group consensus opinion in a collaborative decision process is an inseparable component of many e-commerce activities, such as supplier selection, auctions, scheduling, contracting, etc. [9, 12, 13]. Collaborative decision with larger group consensus requirements is the evolution of the group decision. A consider-

X. Liu (✉)
School of management, Shanghai University, 200444 Shanghai, China
e-mail: lxhelp@sina.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_115

able amount of work on consensus for group decision is available in the literature [1, 2, 3, 5, 9]. They focused on consensus that is simply mathematically derived as opposed to a consensus that is derived from human interaction. However, the collaborative B2B e-commerce implies larger groups of buyers and suppliers in the collaborative decision process, and buyers and suppliers do not form their opinions independently from one another but interactively with other network members [7, 8, 10, 11, 12]. Thus, it is necessary to modify the present vision on current existing models towards the collaborative decision.

We shall establish a social-network-based intelligent collaborative decision model with larger group consensus requirements for B2B e-commerce. It is remarkable that the proposed social-network-based intelligent collaborative decision model is able to allow a certain degree of automation on the consensus-reaching process and integrate opinion leaders and the determination of their degree of influence in a network to consensus measurement by means of applying the theory of social-network analysis to identify the opinion leader and obtain the weigh vector of buyers and suppliers. After that, using the model can decrease the collaborative decision time and maximize the user's utility.

The rest of this chapter is organized as follows. Section 115.2 interprets the social-network-based intelligent collaborative decision model in detail. Section 115.3 demonstrates a detailed implementation process of the social-network-based collaborative decision model, its validity, and applicability. Section 115.4 concludes the chapter.

115.2 The Proposed Model

The proposed social-network-based collaborative decision model includes a 4-phase, which is elaborated step by step in the following:

The first stage Set up collaborative decision social network of a point in time, shown as follows:

Step 1: The suppliers provide the information about the products on an e-commerce platform by the supplier agent. The buyers filtrate, search for products from the e-commerce platform, and get the suppliers' information response to the user by the buyer agent; then generate the collaborative decision social network of a point in time according to the search result, including a group of buyers $B = \{b1, b2, \ldots bt\}(t \geq 1)$ and suppliers $S = \{s1, s2, \ldots sq\}(q \geq 1)$.

The second stage Form collaborative decision problem including three steps, as follows:

Step 2: Select collaborative decision members. The moderator selects collaborative decision members including a group of buyers $B = \{b1, b2, \ldots bt\}(t \geq 1)$ and suppliers $S = \{s1, s2, \ldots sq\}(q \geq 1)$, and then the buyer agents and supplier agents receive proposals from the moderator agent to participate in the collaborative decision problem. Each agent will decide whether it can participate or not (buyer

agent and supplier agent's decision is made by their corresponding human buyer and supplier), and then uses role functions to identify opinion leaders and obtain the weigh vector of buyers and suppliers $\lambda = (\lambda_{b1}, \lambda_{b2}, \ldots, \lambda_{bt}, \lambda_{s1}, \lambda_{s2}, \ldots, \lambda_{sq})^T$, where $\lambda_{bk} \geq 0, k = 1, 2, \ldots, t, \lambda_{sx} \geq 0, x = 1, 2, \ldots, q$ and $\sum_{k=1}^{t} \lambda_{bk} + \sum_{x=1}^{q} \lambda_{sx} = 1$.

Step 3: List the product alternatives. The moderator agent lists the product alternatives $P = \{p_1, p_2, \ldots p_n\}$ and identifies the set of attributes $U = \{u_1, u_2, \ldots u_m\}$ and their types (benefit or cost), and decides the relative weights of the $U = \{u_1, u_2, \ldots u_m\}$ attributes.

Step 4: Let CYCLE be the number of iterative times, and then specify MAXCYCLE, the maximal number of iterative times to avoid the delayed convergence of collective solutions and endless rounds of iterative times; set CYCLE = 0, a cycle counter; specify the threshold of acceptable similarity δ, let $\delta = \delta^*$ be the deadline of acceptable similarity and let $0 < \omega < 1$.

The third stage Automating and supporting collaborative decision consensus-reaching processes include six steps, shown as follows:

Step 5: The moderator agent collects and normalizes the buyer and supplier preferences.

Let $D_{bk} = (d_{ijbk})_{m \times n} (k = 1, 2, \ldots, t)$ and $D_{sx} = (d_{ijsx})_{m \times n} (x = 1, 2, \ldots, q)$ be the buyer decision matrix and the supplier decision matrix, where d_{ijbk}, d_{ijsx} are the attribute values, given by the buyer $b_k \in B$ and the supplier $s_x \in S$, for the product $p_j \in P$ with respect to the attribute $u_i \in U$. The moderator agent collects $D_{bk} = (d_{ijbk})_{m \times n}$, $D_{sx} = (d_{ijsx})_{m \times n}$ and then constructs the buyer $bk \in B$ and the supplier $sx \in S$ normalized decision matrices, shown below as follows: $G_{sx} = (g_{ijsx})_{m \times n}$ and $G_{bk} = (g_{ijbk})_{m \times n}$.

Step 6: Decide whether the collaborative decision consensus is reached. Based on the additive weighted aggregation (AWA) operator, to fuse the normalized decision matrices $G_{bk} = (g_{ijbk})_{m \times n} (k = 1, 2, \ldots, t)$ and $G_{sx} = (g_{ijsx})_{m \times n} (x = 1, 2, \ldots, q)$ into a collaborative decision matrix $G = (g_{ij})_{m \times n}$. For convenience, let $G_{sx}^{(0)} = (g_{ijsx}^{(0)})_{m \times n} = G_{sx} = (g_{ijsx})_{m \times n}$, $G_{bk}^{(0)} = (g_{ijbk}^{(0)})_{m \times n} = G_{bk} = (g_{ijbk})_{m \times n}$, $G^{(0)} = (g_{ij}^{(0)})_{m \times n} = G = (g_{ij})_{m \times n}$.

Let $G_{\text{global}} = (g_{ij(\text{global})})_{m \times n}$ be the global opinion leader decision matrix, $G_{\text{local}} = (g_{ij(\text{local})})_{m \times n}$ be the local opinion leader decision matrix, and $G_{\text{intermediarie}} = (g_{ij(\text{intermediarie})})_{m \times n}$ be the intermediary decision matrix, where $G_{\text{global}} \subset G_{bk} \bigcup G_{sx}$, global $\in \{bk, sx\}$, $G_{\text{local}} \subset G_{bk} \bigcup G_{sq}$, local $\in \{bk, sx\}$, $G_{\text{intermediarie}} \subset G_{bk} \bigcup G_{sq}$, intermediarie $\in \{bk, sx\}$. The illustrated collaborative decision consensus evaluation rule such as if the global opinion leaders and the local opinion leaders and intermediary influence are high (closeness centrality > 0.5, degree centrality > 0.5, betweenness centrality > 0.5), then calculate the similarity degree between $G_{\text{global}}, G_{\text{local}}, G_{\text{intermediarie}}$, and G shown as follows:

$$L(G_{\text{global}}^{(\text{CYCLE})}, G^{(\text{CYCLE})}) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left| g_{ij(\text{global})}^{(\text{CYCLE})} - g_{ij}^{(\text{CYCLE})} \right| \tag{115.1}$$

$$L(G_{\text{local}}^{(\text{CYCLE})}, G^{(\text{CYCLE})}) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left|g_{ij(\text{local})}^{(\text{CYCLE})} - g_{ij}^{(\text{CYCLE})}\right| \tag{115.2}$$

$$L(G_{\text{intermediarie}}^{(\text{CYCLE})}, G^{(\text{CYCLE})}) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left|g_{ij(\text{intermediarie})}^{(\text{CYCLE})} - g_{ij}^{(\text{CYCLE})}\right| \tag{115.3}$$

If $L(G_{\text{global}}^{(\text{CYCLE})}, G^{(\text{CYCLE})}) \le \delta^*$, $L(G_{\text{local}}^{(\text{CYCLE})}, G^{(\text{CYCLE})}) \le \delta^*$ and $L(G_{\text{intermediarie}}^{(\text{CYCLE})}, G^{(\text{CYCLE})}) \le \delta^*$, then go to Step 8, otherwise, go to Step 7.

Step 7: If CYCLE > MAXCYCLE then output a conclusion of no consensus, go to Step1. Otherwise, let $G_{bk}^{(\text{CYCLE}+1)} = (g_{ijbk}^{(\text{CYCLE}+1)})_{m\times n}, G_{sx}^{(\text{CYCLE}+1)} = (g_{ijsx}^{(\text{CYCLE}+1)})_{m\times n}$, $G^{(\text{CYCLE}+1)} = (g_{ij}^{(\text{CYCLE}+1)})_{m\times n}$, where

$$g_{ijbk}^{(\text{CYCLE}+1)} = \omega g_{ijbk}^{(\text{CYCLE})} + (1-\omega) g_{ij}^{(\text{CYCLE})}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n \tag{115.4}$$

$$g_{ijsx}^{(\text{CYCLE}+1)} = \omega g_{ijsx}^{(\text{CYCLE})} + (1-\omega) g_{ij}^{(\text{CYCLE})}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n \tag{115.5}$$

$$g_{ij}^{(\text{CYCLE}+1)} = \sum_{k=1}^{t} \lambda_{bk}\, g_{ijbk}^{(\text{CYCLE})} + \sum_{x=1}^{q} \lambda_{sx}\, g_{ijsx}^{(\text{CYCLE})}, \; i = 1, 2, \ldots, m, \; j = 1, 2, \ldots, n \tag{115.6}$$

where $\sum_{k=1}^{t} \lambda_{bk} + \sum_{x=1}^{q} \lambda_{sx} = 1$. Then, let CYCLE=CYCLE + 1, go to Step 5.

Step 8: Aggregate the buyers' and suppliers' preferences. Based on the AWA [6] operator, calculate the overall attribute values, shown as follows:

$$g_j^{(\text{CYCLE})} = \sum_{i=1}^{m} w_i g_{ij}^{(\text{CYCLE})}, \; j = 1, 2, \ldots, n \tag{115.7}$$

to aggregate all the elements in the j^{th} column of $G^{(\text{CYCLE})} = (g_{ij}^{(\text{CYCLE})})_{m\times n}$, and then get the overall attribute values $g_j^{(\text{CYCLE})}$ (j=1, 2,…,n) of the product $p_j (j = 1, 2, \ldots, n)$.

Step 9: Rank all the products $p_j (j = 1, 2, \ldots, n)$ in accordance with the values $g_j^{(\text{CYCLE})}$ (j=1, 2,…,n).

Step10: Transfer $G_{bk}^{(\text{CYCLE})}$ and $G_{sx}^{(\text{CYCLE})}$ from the reverse unitarization process into the individual buyer decision matrix $D_{bk}^{(\text{CYCLE})}$ and the individual supplier decision matrix $D_{sx}^{(\text{CYCLE})}$.

The fourth stage The output is shown as follows:

Step 11: End. output CYCLE, $D_{bk}^{(\text{CYCLE})} (k = 1, 2, \ldots, t)$ and the results coming from ranking all the products $p_j (j = 1, 2, \ldots, n)$ to the buyer $bk \in B$ (k=1,2,...,t) by the buyer agents and output CYCLE, $D_{sx}^{(\text{CYCLE})}$ and the results coming from rank-

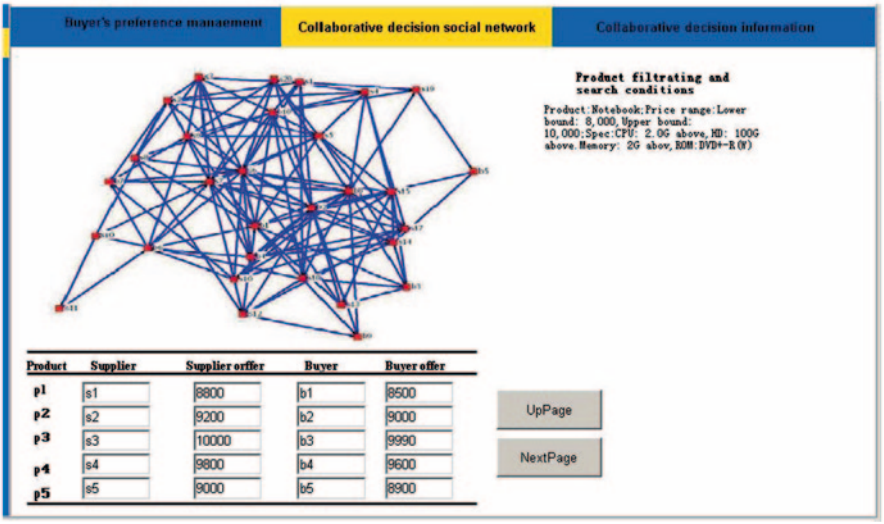


Fig. 115.1 Collaborative decision social network of a point in time

ing all the products $p_j(j=1,2,\ldots,n)$ to the supplier $sx \in S(x=1,2,\ldots,q)$ by the supplier agents.

115.3 Illustrative Example

A self-developed auto-collaborative decision system is developed and used to effectively and efficiently solve the multiple suppliers and buyers dynamic collaborative decision problem. Set up the collaborative decision social network of a point in time (Fig. 115.1).

Form the collaborative decision problem including three steps, such as Fig. 115.2, the buyer $b1$ decision matrix D_{b1} is one of ten buyer decision matrices $D_{bk} = (d_{ijbk})_{5\times5}(k=1,2,\ldots,10)$.

When CYCLE=39, $L(G_{s6}^{(39)}, G^{(39)}) = 0.0707$, $L(G_{b2}^{(39)}, G^{(39)}) = 0.0913$. Since $L(G_{s6}^{(39)}, G^{(39)}) < 0.10$, $L(G_{b2}^{(39)}, G^{(39)}) < 0.10$, the global opinion leader s6 decision matrix $G_{s6}^{(39)}$, the local opinion leader and intermediary b2 decision matrix $G_{b2}^{(39)}$, and the normalized collaborative decision matrix $G^{(39)} = (g_{ij}^{(39)})_{5\times5}$ are of acceptable similarity. Figure 115.3 is the buyer $b1$ decision matrix $D_{b1}^{(39)}$, where the predefined consensus is reached.

In order to evaluate the proposed model performance, let us suppose that a buyer agent asks the user to login to the system and input the same product-filtrating and search conditions (Table 115.1) in seven points in time; and then generate the collaborative decision social networks of seven points in time according to the search

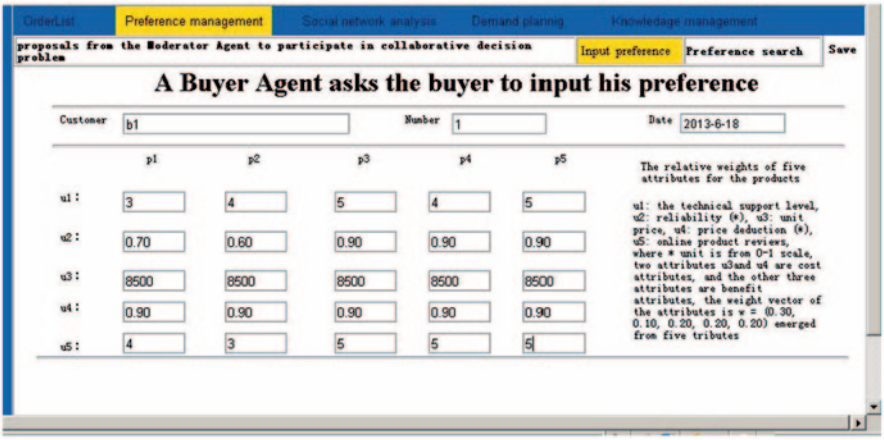


Fig. 115.2 The buyer $b1$ decision matrix D_{b1}

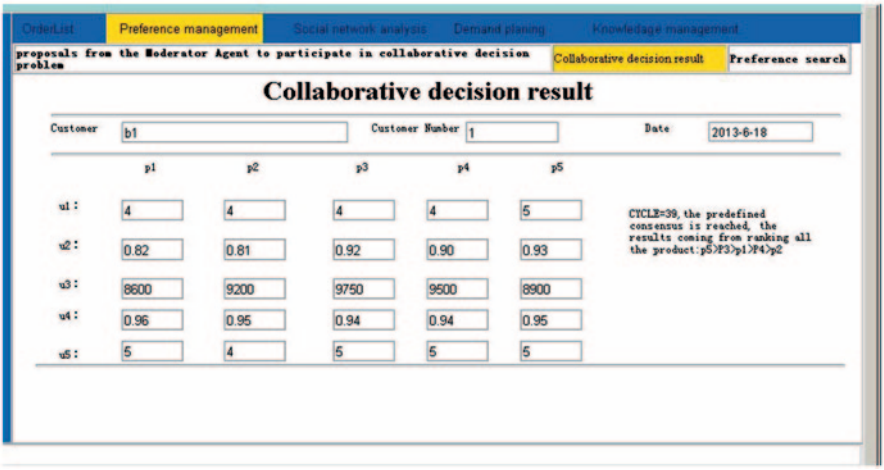


Fig. 115.3 The buyer $b1$ decision matrix $D_{b1}^{(39)}$, where the predefined consensus is reached

result, including a group of buyers B^j and suppliers S^j. Results of the collaborative decision of seven points in time in collaborative decision social network, which are summarized in Table 115.1.

The results show that the consensus is achieved in 39 iterative times in the first point in time. The number of manual supervision and iterative times tend to decrease as the collaborative decision number develops. A low number of suppliers and buyers' human supervision are necessary, which contributes to saving much time and cost in a long term. No supplier and buyer are required to supervise their preferences at each round of the consensus-reaching process. Such supervisions are only necessitated at the beginning, in the first point in time due to the fact that all suppliers and buyers independently give their initial decision matrices.

Table 115.1 Results of the collaborative decision of seven points in time

NO.	Points in time	Number of groups	Manual supervision	Cycle
1	2013-6-18 (first point)	$B^1 = \{b1, b2, \ldots b10\}\ S^1 = \{s1, s2, \ldots s20\}$	30	39
2	2013-6-19 (second point)	$B^2 = \{b1, b2, \ldots b5\}\ S^2 = \{s1, s2, \ldots s15\}$	0	20
3	2013-6-20 (third point)	$B^3 = \{b1, b2, \ldots b9\}\ S^3 = \{s1, s2, \ldots s8\}$	0	16
4	2013-6-21 (fourth point)	$B^4 = \{b1, b2, \ldots b9\}\ S^4 = \{s1, s2, \ldots s6\}$	0	15
5	2013-6-22 (fifth point)	$B^5 = \{b1, b2, \ldots b9\}\ S^5 = \{s1, s2, \ldots s8\}$	0	14
6	2013-6-23 (sixth point)	$B^6 = \{b1, b2, \ldots b9\}\ S^6 = \{s1, s2, \ldots s16\}$	0	12
7	2013-6-24 (seventh point)	$B^7 = \{b1, b2, \ldots b5\}\ S^7 = \{s1, s2, \ldots s19\}$	0	10

115.4 Conclusion

In this chapter, we present a social-network-based intelligent collaborative decision model with large-group consensus requirements for B2B e-commerce-deploying intelligent agents to facilitate the autonomous and automatic collaborative decision for on-line buying and selling while providing a fast response to a group of suppliers and buyers who compete to provide and buy their products. In addition, we have implemented an intelligent collaborative decision social-network-based consensus support system upon such a model. The prominent characteristic model is that it can automatically modify the diverging individual opinions of the buyer and supplier so as to reach consensus among large groups of buyers and suppliers' opinions and actualize the automated collaborative decision-making process. Due to its capacity to deal with larger groups of buyers and suppliers in the real world, we aim to apply the system to the entire supply chain of individuals.

Acknowledgements This work was partially supported by funding from the Humanities and Social Research Program Foundation of the Ministry of Education of the People's Republic of China (11YJA630068).

References

1. Boroushaki S, Malczewski J. Measuring consensus for collaborative decision- making: a GIS-based approach [J]. Environ Urban Syst. 2012;34(3):322–332.
2. Cabrerizo FJ, Moreno M, Perez IJ, Herrera-Viedma E. Analyzing consensus approaches in fuzzy group decision making: advantages and drawbacks [J]. Soft Comput. 2010;14(5):451–63.

3. Cabrerizo FJ, Pérez IJ, Herrera-Viedma E. Managing the consensus in group decision making in an unbalanced fuzzy linguistic context with incomplete information [J]. Knowl-Based Syst. 2010;23(6):169–181.
4. Chen S-M, Lee L-W, Yang S-W, Sheu T-W. Adaptive consensus support model for group decision making systems[J]. Expert Syst Appl. 2012;39(2):12580–12588.
5. Chiclana F, Tapia García JM, del Moral MJ, Herrera-Viedma E. A statistical comparative study of different similarity measures of consensus in group decision making [J]. Inf Sci. 2013;221(5):10–123.
6. Dong YC, Xu YF, Li HY, Feng B. The OWA-based consensus operator under linguistic representation models using position indexes [J]. Eur J Oper Res. 2010;203(2):455–463.
7. Iván P, Rosa M, Luis Martínez R. An attitude-driven web consensus support system for heterogeneous group decision making [J]. Expert Syst Appl. 2013;40(10):139–149.
8. Khorshid S. Soft consensus model based on coincidence between positive and negative ideal degrees of agreement under a group decision-making fuzzy environment [J]. Expert Syst Appl. 2010;37(3):3977–3985.
9. Lau RYK. Towards a web services and an intelligent agent-based collaborative decision system for B2B e-commerce [J]. Electron Commer Res Appl. 2007;6(7):260–73.
10. Lu SYP, Lau HYK, Yiu CKF. A hybrid solution to collaborative decision-making in a decentralized supply-chain [J]. Technol Manage. 2012;29(4):95–111.
11. Petra S, Legner C. B2B integration in global supply chains: an identification of technical integration scenarios [J]. J Strategic Inf Syst. 2013;20(9):250–267
12. Shafiei F, Sundaram D, Piramuthu S. Multi-enterprise collaborative decision support system [J]. Expert Syst Appl. 2012;39(8):7637–7651.
13. Zhao K, Yu X. A case based reasoning approach on supplier selection in petroleum enterprises [J]. Expert Syst Appl. 2011;38(6):6839–6847.

Chapter 116
Toward the Association Rules of Meteorological Data Mining Based on Cloud Computing

Limin Liu, Jie Lv, Zhiqiang Ma, Jianxiong Wan and Ma Jingjing

Abstract Data mining based on association rules (AR) can find out the potential relationship among the meteorological data to forecast the weather conditions; however, it is difficult for traditional AR algorithms to meet the requirement of massive meteorological data mining. To address this issue, we propose a frequent pattern (FP)-tree algorithm based on MapReduce to forecast weather in a cloud computing system. We conduct extensive experiment with real meteorological data to evaluate the performance of our algorithm.

Keywords Association rules · FP-tree · Meteorological data · MapReduce

116.1 Introduction

The meteorological data contains plenty of hidden meteorological laws. Association rule mining (ARM) can find out the potential relationship between the data, which can be used to effectively forecast the weather. The efficiency, adaptability, and usability bottlenecks are available in mining the massive meteorological data with

L. Liu (✉) · J. Lv · Z. Ma · J. Wan · M. Jingjing
College of Information Engineering, Inner Mongolia University of Technology, 010080 Hohhot, China
e-mail: liulimin789@126.com

J. Lv
e-mail: lvjiehappylife@126.com

Z. Ma
e-mail: mzq_bim@163.com

J. Wan
e-mail: wanjx@126.com

M. Jingjing
e-mail: 117967738@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_116

the traditional ARM algorithms [1]. The idea behind cloud computing is to allocate resource on demand dynamically and process massive data at low cost. With this technology, people can process massive amount of data in a short time [2].

There have been some studies seeking to make use of a few machine-learning methods to derive forecasting rules from observational meteorological data; however, those methods fall short of deriving any explicit and direct association relationships by data mining for weather prediction [3].

The traditional ARM algorithm—a priori algorithm—encounters many difficulties in dealing with massive data: It needs to repeatedly scan the database to get all frequent sets; it has high I/O overhead for scanning; it spends so much time in dealing with a particularly large number of candidate sets.

Considering that many meteorological elements are correlated in nature, in this chapter, we design a frequent pattern (FP)-tree algorithm for mining meteorological data [4]. It only needs to scan the database twice and directly combines the database to produce final FPs. Our algorithm includes two steps: (1) the map step takes an input pair and produces a set of intermediate key/value pairs, (2) the reduce step merges these values to form a possibly smaller set of values. We run the FP-tree algorithm on a real meteorological data set to forecast the weather. Some results show that the running time of the algorithm has decreased obviously by increasing the number of computing nodes.

This chapter proceeds as below: Sect. 2 introduces the MapReduce framework; Sect. 3 introduces the MapReduce-based FP-tree algorithm; Sect. 4 analyzes the experimental results; Sect. 5 concludes the chapter.

116.2 Introduction of MapReduce

MapReduce is a parallel-operation programming model or computing model used for large-scale datasets. It mainly contains two concepts: map and Reduce to complete the map and reduce, respectively. Users specify the computation in terms of a map and a reduce function, and the running parallelizes the computation across large-scale clusters of machines, handle machine failures, and schedule inter-machine communication to make efficient use of the network and disks. The map operation is independent [5]. It meets the demands to operate each element in the independent group and then creates a new group to save intermediate values. The elements of an element set are properly merged by reduction. The reduce only obtains a simple answer, even the higher parallelism of the map [6, 7].

Figure 116.1 is the data flow diagram of the MapReduce framework. Firstly, a large number of data sets are divided into a number of small sub-data sets by the map operation; secondly, a sub-data set is assigned to a node; thirdly, each node processes its sub-data set to produce intermediate results; finally, the reduce operation merges all intermediate results into a single final result.

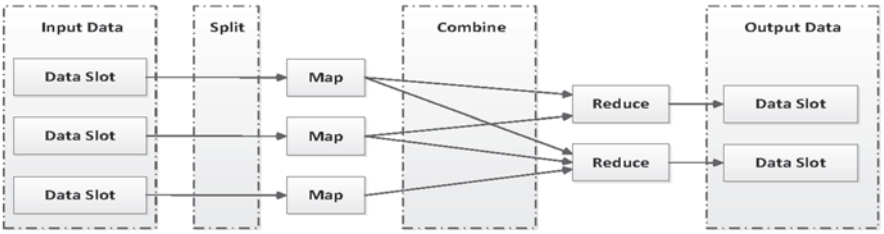


Fig. 116.1 Data flow diagram of MapReduce

116.3 FP-Tree Based on MapReduce

The FP-tree-based algorithm is to divide the database and then distribute each part of the database to nodes or processors for mining, or to distribute the computation loading [7, 8]. During the mining process, the nodes will exchange required transactions with each other. Simultaneously, all the algorithms proposed previously need to scan the local transaction database twice. The FP-tree is a compact representation of all relevant frequency information in a database. Every branch of the FP-tree represents a frequent item set and the nodes along the branches are stored in decreasing order of the frequency of the corresponding items with leaves representing the least frequent items [4, 9].

Our algorithm can divide into two parts, respectively, the computation of MapReduce twice, at last, combination to produce the final FPs. Figure 116.2 is the detailed process of the MapReduce FP-tree algorithm.

The explanations of the MapReduce FP-tree algorithm are as follows:

1. Format the input data according to the Hadoop requirement and store data in the Hadoop distributed file system (HDFS).
2. Scan the transaction database D and divide the data into several data blocks, via a MapReduce process, in which the same attributes gather together and calculate the times of attributes that appear, and then judge whether it is less than the minimum support degree and export the FPs that the attribute is no less than the minimum support degree.
3. Group the FP collections with each group with a unique number; map every frequent item set to the number to formulate the G-list table.
4. According to the G-list table, group the transaction database. As to the mapper of each node, scan the assigned data blocks and process the data items, packet them averagely as much as possible, and then allocate them to the reducer of this group.
5. The reducer of each node processes the assigned data block based on packet, constructing the sub-tree of the FP-tree for data sets, obtaining the conditional pattern bases and conditional FP-tree by iteration. Then get a maximum K FP set.

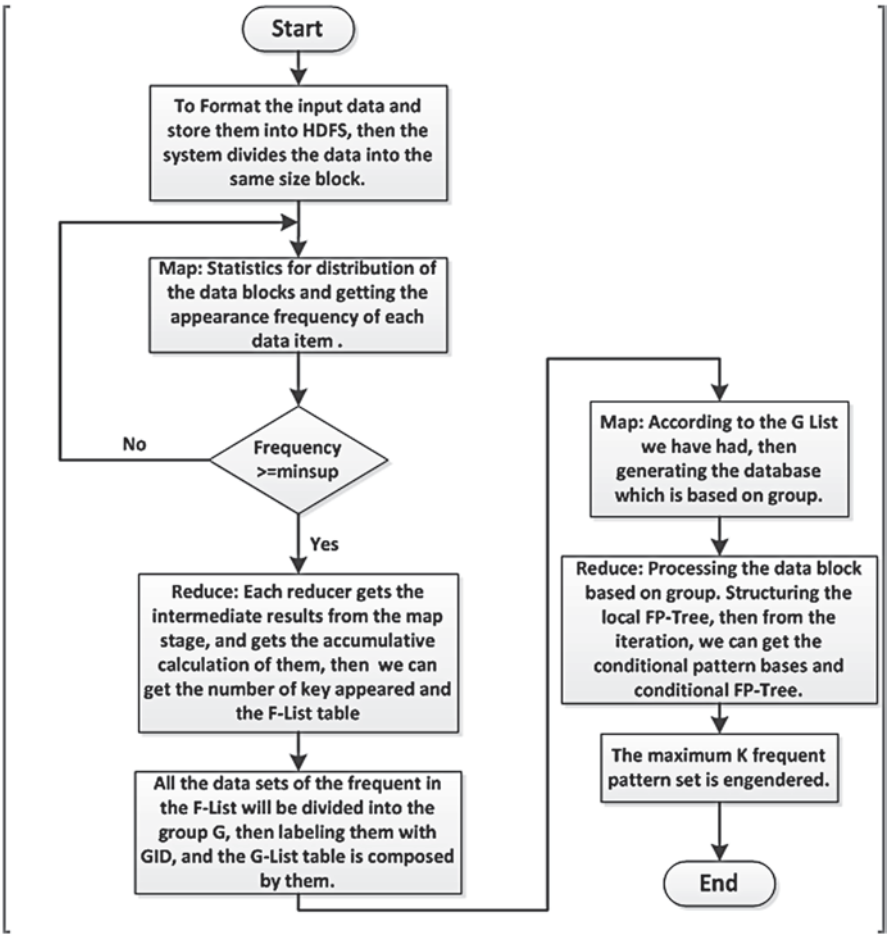


Fig. 116.2 Algorithm flowchart of MapReduce FP-tree

116.4 Experimental Results and Analysis

116.4.1 Mining the Meteorological Data

Experiments use the completely distributed environment in the LAN by using two servers and a disk array to invent PC constructs three nodes cluster. The virtual machine has 4 GB memory, 100 GB hard disk, two 2-core central processing unit (CPU), the Ubuntu 12.04 operating system by using Hadoop 1.0.3 as the test platform, and the Eclipse 3.7 for IDE programming environment.

The data used for predicting experiments are the data set of the China ground international exchange station climatic information provided by China Meteorological Data-Sharing Service Network. The data sets provide daily observation data of 194 basic, reference ground meteorological observation stations and automatic

Table 116.1 Partial data

Station index number	Year	Month	Day	20–20 h precipitation	Average station pressure	Average wind speed	Average temperature	Average vapor pressure	Average relative humidity	Small evaporation rate
53463	1951	2	1	0	8993	38	32766	25	61	28
53463	1951	2	2	0	9015	52	−82	20	64	11
53463	1951	2	3	39	9001	23	−67	31	83	3
53463	1951	2	4	15	9032	12	−134	24	92	17
53463	1951	2	5	0	9055	8	−237	7	78	13
53463	1951	2	6	0	9023	14	−225	9	84	9
53463	1951	2	7	0	8956	0	−189	13	78	11
53463	1951	2	8	57	8901	7	−116	25	92	25
53463	1951	2	9	1	8959	27	−140	17	86	16
53463	1951	2	10	0	8983	5	−225	7	90	8
53463	1951	2	11	0	8980	7	−204	10	83	11
53463	1951	2	12	0	8973	8	−108	24	80	28
53463	1951	2	13	2	8957	13	−92	30	91	12
53463	1951	2	14	0	8974	12	−139	19	93	2
53463	1951	2	15	0	8996	50	−132	21	86	19

reported stations in China between January 1, 1951 and July 31, 2012. The 53,463 (Hohhot station) data include their factor conditions: 20–20-h precipitation, average station pressure, average wind speed, average temperature, average vapor pressure, average relative humidity, and small evaporation rate (Table 116.1).

116.4.2 Prediction of Association Rules in the Meteorological Data

The Boolean association rules process the discretization and kinds of values. They show the relationship between these variables. As to the meteorological association rules data mining, our idea is to discretize the data, converting its properties into Boolean before Boolean ARM (Table 116.2).

Assume that various factors are divided into several grades and marked with a symbol. 20–20-h precipitation uses symbol J, the average wind speed uses symbol F, the average temperature uses symbol W, the average relative humidity uses symbol S, the small evaporation rate uses symbols Z, as shown in Table 116.2.

The ARM using support and confidence cannot determine whether the ARM is meaningful. Based on this problem, it is proposed to calculate the correlation of rules.

Table 116.2 Some symbolized meteorological data records

The class of average wind speed	The class of average temperature	The class of average vapor pressure	The class of average relative humidity	The class of small evaporation rate
F(2)	W(2)	Y(1)	S(3)	Z(1)
F(2)	W(3)	Y(1)	S(3)	Z(1)
F(3)	W(3)	Y(1)	S(2)	Z(2)
F(4)	W(3)	Y(1)	S(2)	Z(2)
F(1)	W(4)	Y(2)	S(2)	Z(3)
F(2)	W(4)	Y(3)	S(4)	Z(1)
F(1)	W(4)	Y(3)	S(3)	Z(2)
F(1)	W(2)	Y(1)	S(3)	Z(1)
F(5)	W(1)	Y(1)	S(2)	Z(1)
F(4)	W(1)	Y(1)	S(2)	Z(1)

The relevancy of V is the than of the possibility of simultaneous events and the possibility of completely independent events. When $V<1$, it indicates negative correlation between events, and this rule is meaningless; when $V=1$, it indicates independence relation between events; when $V>1$, it indicates positive correlation between events and that the rule is meaningful.

As to the data shown in Table 116.2 as an example, mine it with the association rules and calculate the confidence and correlation degree with the results obtained and shown in Table 116.3:

According to the rule table, the following rules can be summed up:

- When the average wind speed is 2 stages, the small evaporation rate is 1 stage, the confidence level can reach 100%.
- When the average vapor pressure is 1 stage, the average relative humidity is 2 stages, the confidence level can reach 57%.
- When the average vapor pressure is 1 stage, the average relative humidity is 3 stages, the confidence level can reach 43%.
- When the average vapor pressure is 1 stage, the average relative humidity is 3 stages, the confidence level of the small evaporation can reach 57%.

Table 116.3 Rule table

Rule number	Rule left	Rule right	Degree of confidence (%)	Correlation
1	Z(1)	F(2)	100	5.56
2	S(2)	Y(1)	57	1.63
3	S(3)	Y(1)	43	1.54
4	Z(1)	Y(1) S(3)	100	5.56
5	Z(1)	S(3)	75	3.125
6	Y(1)	W(3)	100	4.76
7	Z(1)	Y(1)	71	1.69

- When the average relative humidity is 3 stages, the average small evaporation is 1 stage, the confidence level can reach 75 %.
- When the average temperature is 3 stages, the average vapor pressure is 1 stage, the confidence level can reach 100 %.
- When the average vapor pressure is 1 stage, the small evaporation rate is 1 stage, the confidence level can reach 71 %.

According to the obtained rules, it can predict the future event. Among the obtained rules, the occurrence possibility of the left end of the high confidence is also high, which deserves attention.

116.4.3 Performance Evaluation

The data of the performance test use five blocks from the dataset. The size of the blocks is 680 KB, 5.5 G, 11 G, and 22 G. We use these five blocks of data to carry out three groups of experiments.

116.4.3.1 On Single Machine and Distributed Environment

The first group of experiments carry out experimental tests on single machine and distributed environment with the five data sets. The minimum support degree is set to Minsub = 1000. The comparison of the experimental results is shown in Fig. 116.3a:

From Fig. 116.3a we can see that testing the small data sets in distributed environment is more time-consuming than in single-machine environment because it takes some time to initiate the Hadoop calculation and occupies the storage space, which results in lower storage efficiency. When the data size exceeds the ability to run in single-machine environment, or the file size exceeds the default size of the data blocks, the advantage of parallel algorithm will be highlighted.

116.4.3.2 Different Minimum Support Degree on Distributed Environment

The second group of experiments carry out experimental tests on distributed environment using the 5.5-G size of data. When the minimum support degree is 500, 700, 1000, 2000, test the impact of support for distributed computing rate with the experimental results shown in Fig. 116.3b.

From Fig. 116.3b we can see that when the support is increasing, the operating time is reducing because generating frequent item sets relies on the G-list. When the support increases, G-list records will be smaller. Thus, the number of group is smaller at the moment of mining and the computing time of the FP-tree is reducing correspondingly. It is worth nothing that the support larger may impact on the last generation rules. In practical application, we should find the support to enable the computing speed and accuracy to achieve the desired balance.

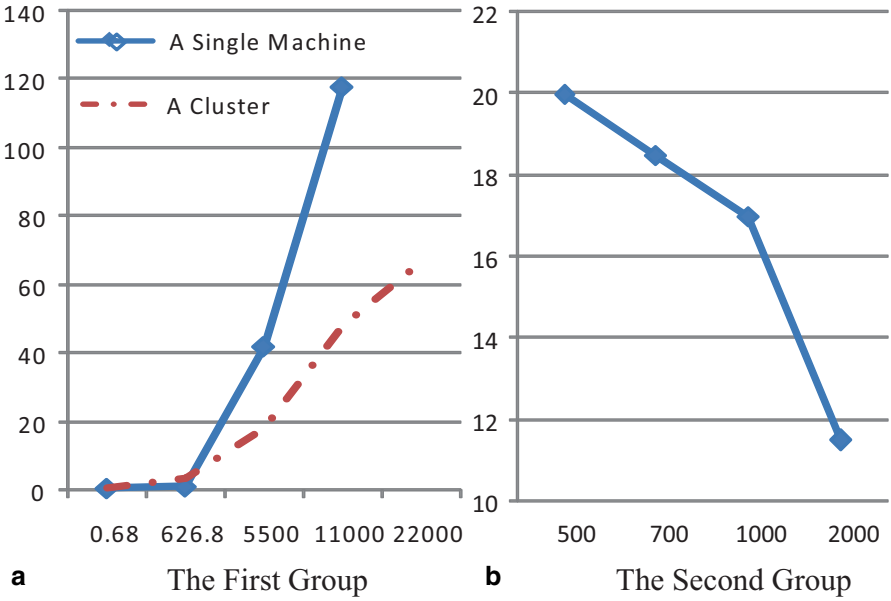


Fig. 116.3 Results and comparison. **a** The first group. **b** The second group

116.4.3.3 On the Number of Different Map Environment

The third group of experiments carry out experimental tests on the number of different map environment with the data of the size of 626.8 MB. By configuring the values of mapred.tasktracker, mapreduce.tasks, maximum of Hadoop, change the number of maps. In the experiments, the number of maps are 2, 4, 6, 8, 10, the running time is 301, 210, 190, 172, 160 s (Fig. 116.4).

The experimental result shows that the running time of the algorithm has decreased obviously by increasing the number of computing nodes. If the computing resources are sufficient, the algorithm can improve the computing speed in a more obvious manner; therefore, as long as appropriately adjusting the number of computing nodes, even the algorithm under the condition of processing data sets on a very large scale, the computing speed will increase correspondingly.

116.5 Conclusion

This chapter introduces a MapReduce implementation of FP-tree algorithm, and discusses its application in terms of meteorological data mining and prediction. With the map operation, the meteorological data-mining problem can be divided into several blocks to produce a set of intermediate key/value pairs. With the reduce operation, it merges together those values to form a possibly small set of values. There are a lot of interesting research issues related to the FP-tree-based mining, including further study, the implementation of constraint-based mining of FPs by

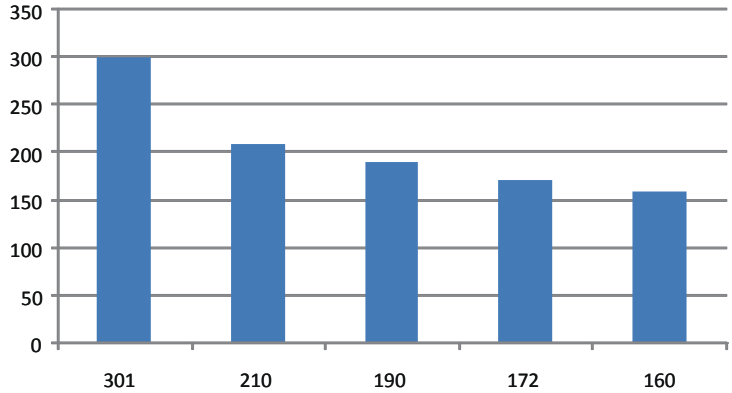


Fig. 116.4 Comparison of the third-group test results

using FP-trees, and the extension of the FP-tree-based mining method for mining sequential patterns, max-patterns, partial periodicity, as well as other interesting FPs.

Acknowledgments This work is supported by No.61363052 Foundation of National Natural Science Foundation of China, No. S20141012812 Inner Mongolia Graduate Research and Innovation Foundation, No. 2014BS0607 Inner Mongolia Provincial Natural Science Foundation, No. NJZY14064 Science Research Project for Inner Mongolia College, and No. S20131012810 Inner Mongolia Graduate Research and Innovation Foundation.

References

1. Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation. ACM SIGMOD Rec. 2000;29(2):1–12.
2. Yong W, Zhe Z, Fang W. A parallel algorithm of association rules based on cloud computing. Proceedings of 2013 8th International ICST Conference on Communications and Networking in China, IEEE; 2013. p. 415–9.
3. Feng L, Dillon T, Liu J. Inter-transactional association rules for multi-dimensional contexts for prediction and their application to studying meteorological data. Data Knowl Eng. 2001;37(1):85–115.
4. Grahne G, Zhu J. Fast algorithms for frequent itemset mining using fp-trees. IEEE Trans Knowl Data Eng. 2005;17(10):1347–62.
5. Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters. Commun ACM. 2008;51(1):107–13.
6. Yu C, Xiao Y, Bo Y. A parallel algorithm for mining frequent itemsets on Hadoop. J Tianjin Univ Technol. 2011;27(1):25–8 (In Chinese).
7. Yang X, Zhen L, Yan F. MapReduce as a programming model for association rules algorithm on Hadoop. Proceedings of 2010 3rd International Conference on Information Sciences and Interaction Sciences, IEEE; 2010. p. 99–102.
8. Zhiyong Z, Chengzhi Y, Tao Y. Research of load balance FP-growth algorithm in parallel. Computer Eng Appl. 2010;46(4):125–6 (In Chinese).
9. Jun G, He S. A fast association rule mining algorithm[J]. Computing Engineering. 2006;32(1):116–8 (In Chinese).

Chapter 117
A Novel In-flight Alignment Algorithm for Airborne Strapdown INS with Assistant of Doppler Radar

Shuang Pan, Xu Han and Nie Yongfang

Abstract In order to improve the precision of Doppler and Strapdown inertial navigation system (SINS) integrated navigation system with moving base, an in-flight alignment method is proposed by making an uniform circular motion of the aircraft according to the phenomena avaliable in the Doppler/SINS navigation system tests. The platform error-angles of the SINS can be estimated quickly and the prediction precise can be improved by this method. Experiment verifies the feasibility and validity of this method. Theoretical analysis shows that the steady state limit of the estimation can be highly reduced by proposed method, and the method is valuable for practical application.

Keywords Airborne SINS · Doppler · In-flight alignment · EKF · Filtering estimation

117.1 Introduction

As limited by the status of INS equipment and the level of theoretical research, the in-flight alignment cannot be implemented effectively just by using the existing INS equipment on airborne which can finish the work of solver on its own [1]. Most scholars tend to believe that the INS resolving capability can be improved with the assistance of some airborne mission equipment, and the attitude misalignment angles can be estimated and compensated effectively by using the auxiliary information as reference as obtained from some established satellite positioning

S. Pan (✉) · N. Yongfang
Naval Submarine Academy, 266041 Qingdao, China
e-mail: n7k9pp@126.com

N. Yongfang
e-mail: gaochao.shd@163.com

X. Han
Naval Aeronautical and Astronautical University, 264001 Yantai, China
e-mail: Hanxu171213@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_117

systems. The in-flight alignment method for airborne SINS with the assistance of its own equipment is also called the in-flight integrated alignment or the stand-alone auxiliary alignment [2–5].

Doppler device is generally considered to be rigidly connected with the body [6–8]. Doppler velocity errors contain inherent errors of the equipment such as scale error, bias and noise, and installation errors such as structural misalignment. Bias and scale factor/sight line axial error are generally considered to be random constants. As the Doppler device is based on the principle of the radar, the noise may be a function of the speed vector and is usually considered as fluctuation noise.

Because of the high speed precision of the Doppler radar speed gun, it is a reasonable idea for us to use the Doppler radar speed gun to assist the moving-base alignment. On this basis, the paper aims at proposing an in-flight emergency alignment method with the Doppler observables and circular motion alignment. The paper uses the EFK filtering algorithm to solve the problem of uncertainty of the measurement noise, and the simulation illustrates that the EFK filtering algorithm can effectively decrease the errors and increase the convergence rate.

117.2 Doppler and Integrated System Error Model

The principle of Doppler radar speed gun is to use the Doppler effect as generated when electromagnetic waves encounter the moving target, that is, the frequency of the echo signal generated by the encounter between the transmitting radar waves (the frequency is f_0) and the moving object is $f_0 \pm f_d$, where f_d is the Doppler frequency shift. If V_r is the radial distance from the target to the radar, the formula for f_d and V_r is $f_d = 2f_0 \times V_r/(c + V_r)$, where c is the velocity of light, and generally, $c >> V_r$. Then it can be obtained that $V_r = f_d \times (\lambda_0/2)$, where $\lambda_0 = c/f_0$ is the wavelength of the emitted electromagnetic wave; therefore, V_r can be obtained as long as f_d (f_0 and c are known) is measured.

The errors of Doppler radar speed gun mainly consist of the following components: (i) the error caused by changes in transmitter frequency (0.0048 %); (ii) the error caused by changes of the velocity of light (0.023 %); (iii) the error caused by measurement methods (0.0002 %); (iv) the error caused by the instability of the clock frequency (0.005 %); (v)the error caused by the changes of the phase which occurs when the signal gets through the receiver, the input amplifier and the band-pass filter (0.02 %); (vi) the measurement error due to the time reference error of the tape recorder when the signal gets through the recorder for playback measurement and when there is no time recorder in the data processing apparatus (0.1 %). When summing up the above several errors, the radar speed error is about 0.1031 %, which meets the accuracy requirements of the navigation [1].

A common error model for Doppler velocity sensor is given by

$$\begin{bmatrix} \delta v^b_{xDOP} \\ \delta v^b_{yDOP} \\ \delta v^b_{zDOP} \end{bmatrix} = \begin{bmatrix} dB_x \\ dB_y \\ dB_z \end{bmatrix} + \begin{bmatrix} dSF_x & dBS_{xy} & dBS_{xz} \\ dBS_{yx} & dSF_y & dBS_{yz} \\ dBS_{zx} & dBS_{zy} & dSF_z \end{bmatrix} \begin{bmatrix} v^b_x \\ v^b_y \\ v^b_z \end{bmatrix} + \begin{bmatrix} dW_x \\ dW_y \\ dW_z \end{bmatrix} \tag{117.1}$$

In this error model, the term related to off-axis and speed are used to denote the misalignment errors or the sight line axial errors. dBS_{yz} and dBS_{zx} denote azimuth and pitch sight line axial errors, respectively. As to the fix-wing airborne application, the product of the error and velocity components v^b_y and v^b_z can usually be ignored because they are very little when compared with the bias of the Doppler device.

The output of the Doppler velocity sensors are combined with the INS velocity, by which a measurement is obtained for the use of processing the Kalman filtering algorithm for airborne. The difference between Doppler velocity and INS velocity is called Doppler bias which can be denoted as follow [1]:

$$\Delta \underline{v}_{DOP} = \underline{v}_{INS} - \underline{v}_{DOP} \tag{117.2}$$

where v_{INS} is the velocity resolved by INS, v_{DOP} is the velocity measured by the radar.

Under normal circumstance, the Carrier-based aircraft in flight requires the pitch angle θ and roll angle γ to be less than 5, and heading angle which is between 0–360° to vary for different headings; therefore, the pitch angle error and the roll angle error have less impact on resolving the aircraft location in latitude and longitude coordinates. The aircraft positioning errors mainly originate from the heading measurement error and the velocity measurement error [3].

117.3 Doppler and Integrated System Error Model

The Earth's rotation angular velocity and the gravitational acceleration are used by INS to conduct autonomous coarse alignment, then the velocity measurement is compared with the INS velocity, and the state vectors are estimated through Kalman filter, the estimated values are used to regulate the INS, and then the initial alignment is completed. Its principle is shown in Fig. 117.1.

The navigation parameters output by the navigation solver and the radar speed gun respectively are fused together through Kalman filter, and the estimated value of the parameter error the filter outputs is used directly to regulate the navigation parameter the SINS outputs. A combination of the output correction and feedback correction will be taken as the correction method. In the course of estimation, the output correction is firstly used and the feedback correction is used as the filter goes stable (that is to say, the estimation errors go stable), then the corrected navigation parameters will be the output of the navigation system. Figure 117.2 shows the navigation procedure of SINS with aid of radar.

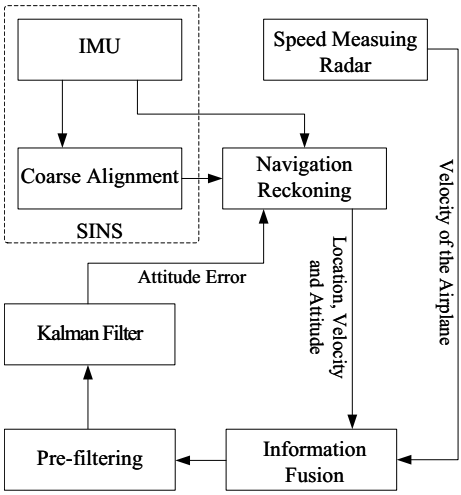


Fig. 117.1 Alignment procedure of SINS with aid of radar

117.3.1 Mathematical Model of Doppler/SINS In-Flight Alignment

1) State equation

In the course of alignment, the body makes a slow rotary motion in the horizontal plane, the coupling of height channel and horizontal channel is weak, moreover, the height channel makes a damping by utilizing such auxiliary altimetry equipment as

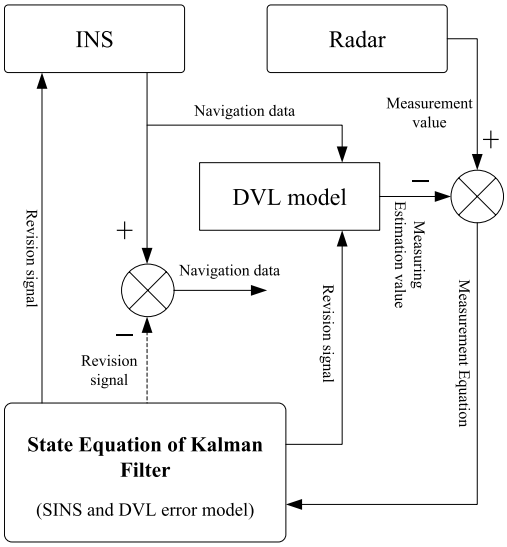


Fig. 117.2 Navigation procedure of SINS with aid of radar

barometric altimeter; therefore, in the approximately horizontal rotation, the height channel parameters change so slightly that the impact of the height channel may be ignored. A local level N-E-D frame is used as the navigation frame, and the initial alignment state vector X is as follows:

$$X = \left[\varphi_E, \varphi_N, \varphi_U, \delta V_x, \delta V_y, \delta L, \delta\lambda, \varepsilon_x, \varepsilon_y, \varepsilon_z, \delta V_d, \delta\Delta, \delta C\right]^T \tag{117.3}$$

where φ_E, φ_N and φ_U are the attitude misalignments; δV_x and δV_y are the east and north velocity errors respectively; $\delta\lambda$ and δL are the longitude and latitude errors respectively; $\varepsilon_x, \varepsilon_y$ and ε_z are the lateral, longitudinal and vertical gyro drifts respectively; δV_d is the velocity offset error,; $\delta\Delta$ is the drift angle error, and δC are the scale factor error.

The system's state equation can be written as below:

$$\begin{aligned}\dot{\varphi}_E &= (\omega_{ie}\sin\phi + \frac{V_x}{R}\tan\phi)\cdot\varphi_N - (\omega_{ie}\cos\phi + \frac{V_x}{R})\cdot\varphi_U \\ &+ \frac{1}{R}\cdot\delta V_y + C_b^n(1,1)\cdot\varepsilon_x + C_b^n(1,2)\cdot\varepsilon_y + C_b^n(1,3)\cdot\varepsilon_z\end{aligned} \tag{117.4}$$

$$\begin{aligned}\dot{\varphi}_N &= -(\omega_{ie}\sin\phi + \frac{V_x}{R}\tan\phi)\cdot\varphi_E - \frac{V_y}{R}\cdot\varphi_U + \frac{1}{R}\cdot\delta V_x \\ &-\omega_{ie}\sin\phi\cdot\delta\phi + C_b^n(2,1)\cdot\varepsilon_x + C_b^n(2,2)\cdot\varepsilon_y + C_b^n(2,3)\cdot\varepsilon_z\end{aligned} \tag{117.5}$$

$$\begin{aligned}\delta\dot{V}_x &= f_U\varphi_E - f_N\varphi_U + \frac{V_x}{R}\tan\phi\cdot\delta V_x \\ &+ \begin{pmatrix} 2\omega_{ie}\sin\phi \\ +\frac{V_x}{R}\tan\phi \end{pmatrix}\cdot\delta V_y + (2\omega_{ie}V_y\cos\phi + \frac{V_xV_y}{R}\sec^2\phi)\delta\phi + \nabla_E\end{aligned} \tag{117.6}$$

$$\begin{aligned}\delta\dot{V}_y &= -f_U\varphi_E + f_E\varphi_U - 2(\omega_{ie}\sin\phi + \frac{V_x}{R}\tan\phi)\cdot\delta V_x \\ &-(2\omega_{ie}V_x\sin\phi + \frac{V_xV_y}{R}\sec^2\phi)\cdot\delta\phi + \nabla_N\end{aligned} \tag{117.7}$$

$$\begin{aligned}\delta\dot{V}_y &= -f_U\varphi_E + f_E\varphi_U - 2(\omega_{ie}\sin\phi + \frac{V_x}{R}\tan\phi)\cdot\delta V_x \\ &-(2\omega_{ie}V_x\sin\phi + \frac{V_xV_y}{R}\sec^2\phi)\cdot\delta\phi + \nabla_N\end{aligned} \tag{117.8}$$

$$\delta\dot{L} = \frac{1}{R}\cdot\delta V_y, \dot{\varepsilon}_x = -\frac{1}{T_g}\cdot\varepsilon_x + \omega_x; \dot{\varepsilon}_y = -\frac{1}{T_g}\cdot\varepsilon_y + \omega_y, \dot{\varepsilon}_z = -\frac{1}{T_g}\cdot\varepsilon_z + \omega_z$$

where f_E, f_N, and f_U are the specific force vectors measured by accelerometer in the navigation frame; V_E, V_N, L and λ are navigation parameters output by the system, ∇_E and ∇_N are the accelerometer random noises which obey a normal distribution with zero mean.

2) Observation equation

In this algorithm, set the observation as the difference between the velocity value resolved by SINS and the velocity value measured by Doppler radar, then system observation vector Z_k is given by

$$Z_k = [\delta V_x - \delta V_{dE}, \delta V_y - \delta V_{dN}]^T \tag{117.9}$$

117.3.2 Filter Design

The above state equation and observation equation can be expressed as follows:

$$\begin{cases} \dot{X} = A(t)X + GW \\ Z_k = H_k X_k + V_k \end{cases} \tag{117.10}$$

where the vector of observation noise is $v = [v_{kE}, v_{kN}]^T$, the observation matrix H is given by

$$H = \begin{bmatrix} -V_y & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -\sin k_d & -V_x & -V_y \\ V_x & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -\cos k_d & -V_y & V_x \end{bmatrix} \tag{117.11}$$

where k_d denotes the consideration for the track direction of drift angle.

The system state equation is continuous and linear, and the observation equation is discrete and linear. The discrete state equation can be written as follows:

$$\begin{cases} X_k = \varphi_{k,k-1} X_{k-1} + \Gamma_{k,k-1} W_{k-1} \\ Z_k = H_k X_k + V_k \end{cases} \tag{117.12}$$

where $\varphi_{k,k-1}$ is the state transition matrix, $\Gamma_{k,k-1}$ is the model noise transition matrix, W_k is the model noise, H_k is the observation matrix, V_k is the observation noise.

The discrete Kalman filtering equations are formulated as follows:

The state step prediction:

$$\hat{X}_{k,k-1} = \varphi_{k,k-1} \hat{X}_{k-1}$$

The state estimation:

$$\hat{X}_k = X_{k/k-1} + K_k(Z_k - H_k\hat{X}_{k/k-1})$$

The filter gain:

$$K_k = P_{k/k-1}H_k^T(H_kP_{k/k-1}H_k^T + R_k)^{-1}$$

The step prediction mean square error:

$$P_{k/k-1} = \varphi_{k,k-1}P_{k-1}\varphi_{k,k-1}^T + Q_{k-1}$$

The estimated mean square error:

$$P_k = (I - K_kH_k)P_{k/k-1}(I - K_kH_k)^T + K_kR_kK_k^T$$

117.4 Simulation and Analysis

117.4.1 Simulation Environment Settings

In the simulation experiment, as the body mainly makes a short time circular motion in the process of alignment, the INS error model previously established can be simplified, and the initial simulation parameters are set as follows:

- The initial position of 30 °N, 120 °E, initial azimuth, pitch and roll angles are chosen as 0 ° respectively, the initial leveling and the azimuth misalignment angle are chosen as 6 arc min and 10 arc min respectively.
- The flight path design: in the previous 50 s, the carrier aircraft keeps climbing up at a speed of 60 m/s, and then keeps a straight flight at a stable speed of 60 m/s in the subsequent 50 s; then, with a body roll angle of 30°, the aircraft keeps doing circling maneuver at a angular velocity of 10°/s in 300 s, then the aircraft keeps a uniform level flight, the total simulation time is 400 s.
- A medium accuracy (1 n mile/s) SINS is chosen, the gyro drift relevant time τ is chosen as 10,800 s, the constant and random drifts of each gyro are chosen as 0.02 (°)/h and 0.01 (°)/h (1σ), respectively, the random constant rms is chosen as 4.8×10^{-8} rad, the accelerometer bias relevant time is chosen as 10,800 s, the constant and random biases of each accelerometer are chosen as $10^{-4}g$ and $5\times10^{-5}g$ (1σ) respectively, the constant RMS is chosen as 4.8×10^{-8} rad.

117.4.2 Simulation Results and Analysis

The carrier aircraft makes some kinds of maneuvering flight such as climb flight, speed flight, level flight and circling flight; the total simulation time is 400 s; the alignment is performed in the circular motion, during which the resolved data of the Doppler/SINS is fused with the position and velocity measured by GPS; then the work of alignment is completed. Figures 117.3, 117.4, 117.5 show the position error, velocity error and error of the misalignment angle estimation in initial alignment.

The simulation results shown in Figs. 117.3, 117.4, 117.5 illustrate that the alignment algorithm for airborne SINS on moving base with the assistant of Doppler radar can get stable filtering errors through the filtering solver in less than 30 s. The attitude misalignment angle estimation errors are as follows: $\Delta\varphi_E = 25.7'', \Delta\varphi_N = 26.4'', \Delta\varphi_U = 3.24'$; and the velocity errors are as follows: $\Delta V_E = 0.46m/s, \Delta V_N = 0.39m/s$; in addition, the positioning errors of SINS assisted by the GPS in position matching are as follows: $\Delta L = 12.7m, \Delta\lambda = 15.3m$.

The simulation results illustrate that the algorithm can obtain considerable alignment accuracy in a short time. In addition, the algorithm can effectively solve the problem of the uncertainty of measurement noise and increase the convergence rate. Compared with the theory in some relevant literatures that the steady-state errors of the 3 misalignment angles are the ultimate errors and can not be eliminated; and the

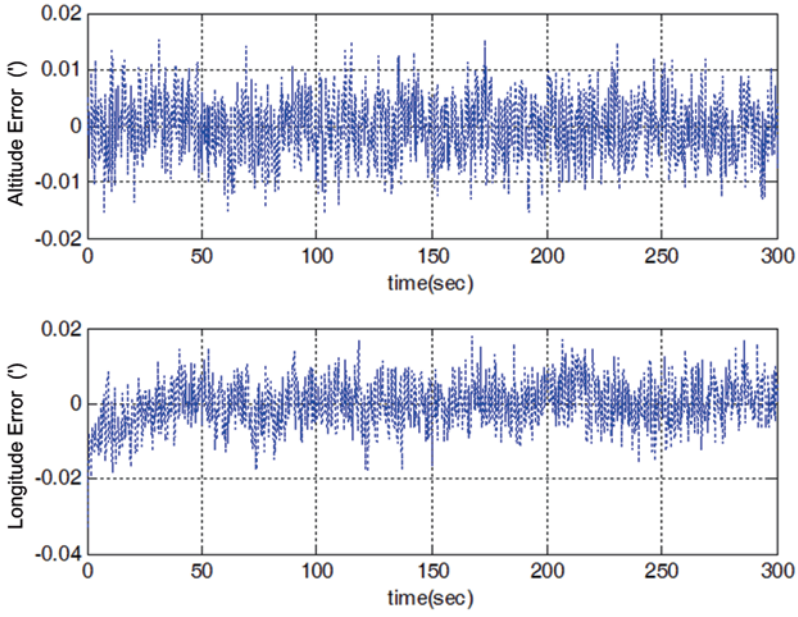


Fig. 117.3 Localization error in alignment

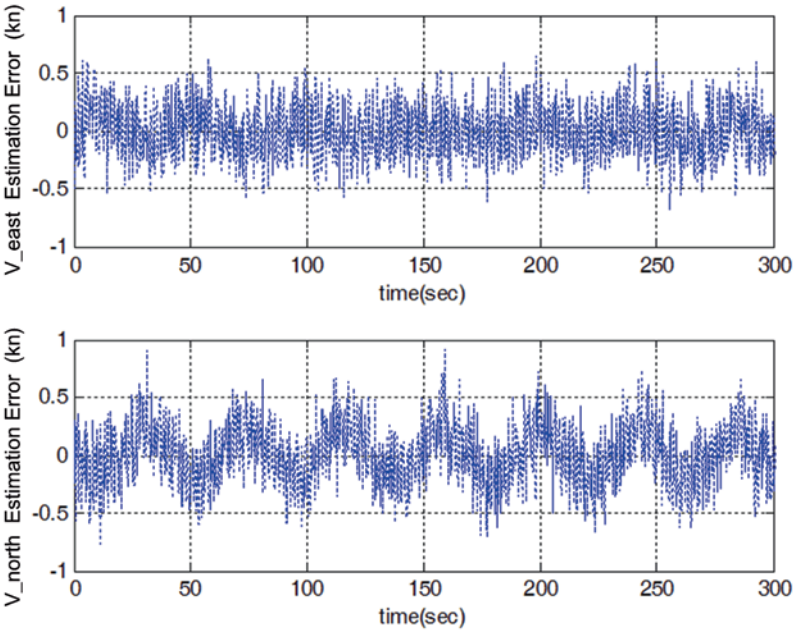


Fig. 117.4 Velocity error of airplane in alignment

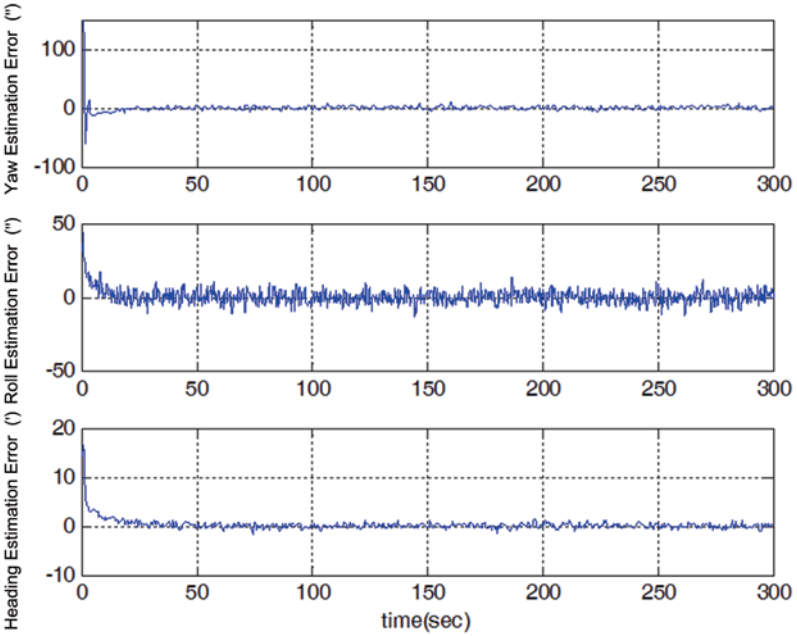


Fig. 117.5 Misalignment error of SINS in alignment

alignment algorithm can be regarded as an effective method to suppress the random errors of gyro and accelerometer. Furthermore, for the sake of obtaining the accurate three-dimensional velocity information, it is very important to get the optimal radar output by integrating variety of factors.

117.5 Conclusion

As to the issue that the respond time is affected by the ground static alignment of SINS/GPS integrated navigation system, the paper proposes an in-flight emergency alignment method with the Doppler observables and the circular motion alignment.

The system state equation and observation equation were founded, and the car test verified the method to be feasible and effective. The method can greatly improve the speed and precision of the alignment and be regarded as an effective method to suppress the random errors of gyro and accelerometer. It may be noted that it would be an effective way of engineering implementation by combining a high-precision single-axis rate turntable with SINS into one system.

References

1. Muhammad JA, Ushaq M. A consistent and robust Kalman filter design for in-motion alignment of inertial navigation system. Measurement. 2008;10(2):1–6.
2. Wang X, Guo L. An intelligentized and fast calibration method of SINS on moving base for planed missiles. Aerosp Sci Technol. 2009;4(4):1–8.
3. Wang X. Fast alignment and calibration algorithms for inertial navigation system. Aerosp Sci Technol. 2009;4(2):1–6.
4. Chan GP, Kim K. UKF based In-flight alignment using low cost IMU. AIAA GNCCE. 2006;08(9):1–12.
5. Wang X, Gong D, Xu L, Shao X. Laser radar based relative navigation using improved adaptive Huber filter. Acta Astronautica. 2011;68(8):1872–80.
6. Hajiyev C. Fault tolerant integrated radar/inertial altimeter based on nonlinear robust adaptive Kalman filter. Aerosp Sci Technol. 2012;17(7):40–9.
7. Kim DY, Jeon M. Data fusion of radar and image measurements for multi-object tracking via Kalman filtering. Inform Sci. 2014;27(6):43–9.
8. Wang Q. An adaptive fusion algorithm based on ANFIS for radar/infrared system. Expert Syst Appl. 2009;36(3):111–20.

Chapter 118
Design of Router for Network on Chip Based on Network Coding

Xisong Fu, Huaxi Gu, Ke Chen, Rong Gao and Jian Zhu

Abstract Network on Chip (NoC) has significantly solved the problem caused by the bus architecture of System on Chip (SoC). The multicast communication will become an increasingly important factor in future NoC-based system. Network coding (NC) can help to improve the performance of NoC. In this paper, we design a router architecture based on network coding. The design principle and working process of this new router, especially the coding and decoding modules are discussed in details. Finally, a simulation model of the router is established.

Keywords Network on chip · Network coding · Router

118.1 Introduction

Network on Chip (NoC) is proposed to solve the communication challenges of the System on Chip (SoC). Multicast becomes one of the most important communication styles, as more and more on-chip applications appear. In the distributed shared cache system, multicast has taken 12.4 % of the total account of communication [1]. Unicast NoC cannot satisfy the multicast communication very well. Many multicast mechanisms, such as tree based [2] or path based [3], have been presented to optimize the performance of NoC; however, NoC cannot achieve the maximum value of max-flow/min-cut theorem [4]. This will be the bottleneck of the throughput. Network coding (NC) is an efficient technique to solve this problem [4]. With this technique, two or more packets received and stored previously are coded together to generate one outgoing packet in the middle nodes, and at the destination node, the

X. Fu (✉) · H. Gu · K. Chen · R. Gao
State Key Laboratory of ISN, Shenzhen CU-Xidian joint center,
Xidian University, 710071 Xi'an, China
e-mail: afuxiaosong@163.com

J. Zhu
Key Laboratory of Network Coding Key Technology and Application, Shenzhen Research Institute, The Chinese University of Hong Kong, 518057 Shenzhen, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_118

coded packet can be decoded; therefore, NC increases the capacity of the link and also eliminates the competition of the packets, which are coded together.

NC is introduced to NoC recently [5–8]. The router architecture working on the centralized control mode [6] and routing algorithm based on bus [9] are designed; however, the architecture of the NC module is ignored. In this paper, we design a new architecture of NoC router based on NC, including input (IPT) module, NC module, Switch Allocation (SA) module, crossbar (Xbar) module, and output (OPT) module. This paper mainly discusses the design of the NC module. The new router works on the distributed control mode, and it makes the NC operation more flexible and enhances the scalability of NoC.

The rest of this paper is organized as below: in Sect. 118.2, we introduce the architecture and the packet format of the NoC router based on NC. The detailed architecture of NC module is provided in Sect. 118.3. The evaluation and simulation results are presented in Sect. 118.4. Finally, the paper is concluded in Sect. 118.5.

118.2 NC-Based Router Architecture

The structure of the new proposed router is shown in Fig. 118.1a. IPT module containsrouting computation, timing, and input buffer unit. Coding and decoding module belong to NC module. SA module is used to allocate Xbar and implemented by the simple round robin arbitrators. The allocation may involve competition not only for the output port of Xbar, but also competition with other virtual channels in the same IPT module for the input port of the Xbar. Moreover, the packets from NC module is handled no differently than any other packets in both SA and Xbar module. Xbar interconnects IPT and OPT module, and connects the NC module and OPT module, either. Xbar and OPT module have dedicated ports for the NC module to decrease the blocking probability. The Virtual Channel Allocation (VA) unit is located in the OPT module.

Each packet travels through the router in the following five stages, Routing Computation (RC), NC, SA, Switch Transmission (ST), and Output. When a packet is injected into the router, it is firstly buffered and the out port is calculated in the RC stage at the same time. Information extracted in RC stage activates the coding or the decoding operation of NC stage. This stage produces the coded packets. The coded and un-coded packets get the permission of Xbar during SA stage. The permission makes the packets to finish the ST stage and get new VC of the next node on the last stage. During this stage, the destination address is updated.

To support the network coding function, a new packet format, referencing to Fig. 118.1b, is introduced for this new router. There is at least one head flit in each packet. The UI in the head flit comprises the source address and the packet number. The packet number is a sequence number created by IP core or NC module for the new packet. Each packet owns a unique UI, and this assists the decoding module to find out the decoding packet. The decoding number named as DN_1 or DN_2 is derived from the UI whose packet joins in the coding operation. DA is the destination

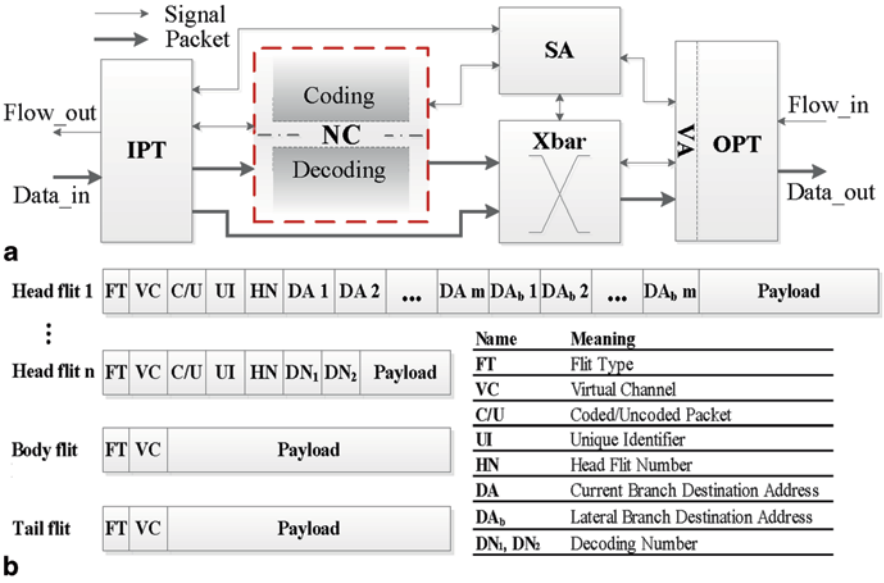


Fig. 118.1 **a** Router structure and **b** Packet format

address for the current branch packet, and DA_b for lateral branch packet. Each bit in the DA and DA_b represents a certain node in the network.

118.3 Network Coding Module

NC module contains the coding module and the decoding module. The two modules provide the coding function and the decoding function.

118.3.1 Coding Module

The structure of coding module is presented in Fig. 118.2. As the new router works on the distributed control mode, the Coding Condition Judgment/Arbitration (CJA) module containing coding judgment unit and arbitration unit is necessitated to decide which packet to join in the coding operation. The Coding Operation (CO) module and the Create Head Flit (CH) module produce the coded packets cooperatively; in addition, the Reorganized Packets (RP) module assembles the coded packet. Coding Buffer (CB) is essential to store the flit coming from CO module. As the coded packet is a new one, UI module is built to create the UI for coded packets. The coding module cannot work smoothly without the Management (MA-c) module. MA-c module includes Flow Control (FC-c), Buffer Allocation (BA-c), and Coding Start/End (CSE) module.

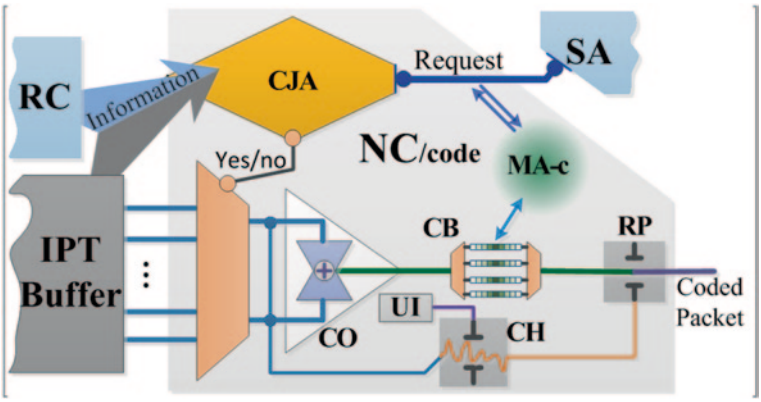


Fig. 118.2 Coding module structure

When a new packet comes into the router, IPT module extracts the information in head flit and calculates the out port. Both of them will be sent to the coding module to judge whether the packet satisfies the coding condition or not.

To save hardware resource, such as buffer and logic unit, two-stage arbitration with Round Robin is used to select two packets to join in the coding operation; then, the BA-c module will allocate a free buffer for the coded packet, which is generated by the two packets.

When the above procedure is ready, packets will be read into the coding module to finish the coding operation. The reading action needs to be controlled by the FC-c module, which manages both input buffer and coding buffer, and makes two packets join in the coding operation synchronously.

As all the flits from the same packet are added into the coding operation, the flits, which are outputted from the CO module, will lose the routing information and turn to un-head flits of coded packet; therefore, a new head flit of coded packet has been created from the CH module. The head flit does not inject into the coding buffer. In this way, the working time can be saved and it is beneficial for the control operation. When the reading signal from SA module comes, the RP module assembles the coded packet. The coded packet routes like the un-coded packet after it outputs from the NC module.

When the packet is not able to passes through CJA or BA module, it will skip the coding operation.

118.3.2 Decoding Module

The structure of decoding module is shown in Fig. 118.3, including Find Decoding Packet (FDP) module, Delete Head Flit (DH) module, Decoding Operation (DO) module, Decoding Buffer (DB), and Management (MA-d) module. MA-d module contains Flow Control (FC-d), Buffer Allocation (BA-d), and Decoding Start/End (DSE) module.

When a coded packet arrives at the destination, it will be stored in the input buffer as an un-coded packet firstly. By the time routing computation unit finds a coded packet has arrived at the destination, it will send a decoding signal to the decoding module.

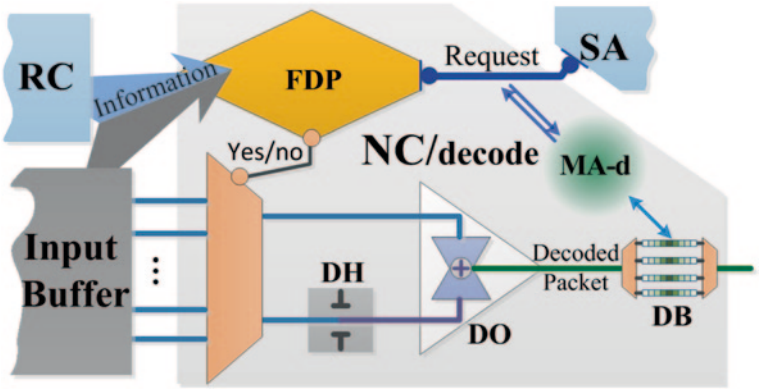


Fig. 118.3 Decoding module structure

When the FDP module gets the decoding signal, it tries to find the decoding packet from the IPT module. Two free decoding buffers, allocated by BA-d module, are necessitated for the decoding operation to store the decoding packet and the decoded packet. The packets will be read into decoding module to decode the coded packet when the above is ready. The reading signal is similar to the reading signal in the coding module.

A shield signal will be created to stop the timing function in the IPT module and the reading of decoding packets from SA module when the decoding packet is found, in case the decoding packet is sent out of this node before the decoding operation ends. As the router is working in the distribution mode, they cannot inform each other which packet is the decoding packet, and only the coded packet knows which one is the decoding packet in the network; therefore, when un-coded packet reaches the destination node, it waits several working cycles for the coded packet to confirm that it is a decoding packet or not. The timing function is used to count the number of the working cycle for the un-coded packet. When the timeout occurs, the packet is ejected from the network.

The coded packet deletes the head flit in the DH module before the decoding operation. As the head flit is added into the coded packet after the coding operation, it cannot join in the decoding operation, either. After the decoding operation, the packet is stored in the DB, and the switch allocation request signal is created in the meantime. During the decoding operation, the decoded packet's flit type should be reorganized because the flit type has been reorganized in coding operation.

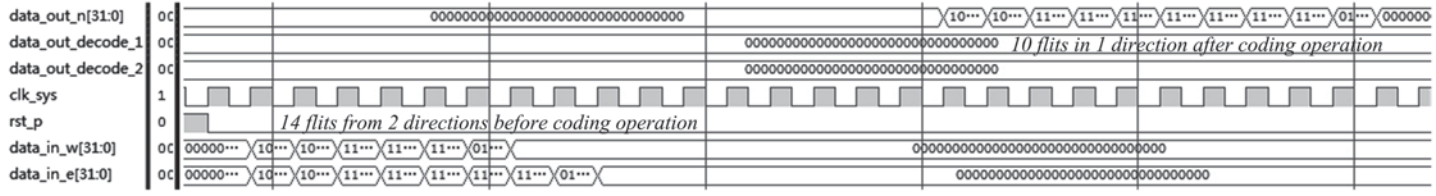
118.4 Result

In this section, we will show our simulation model of this new router. This model has been built with Verilog HDL language with the simulation conducted with the ISE design suite. Our configuration of router has four virtual channels in each port. Each of these channels has a 16 depth FIFO and the data width is 32 bits.

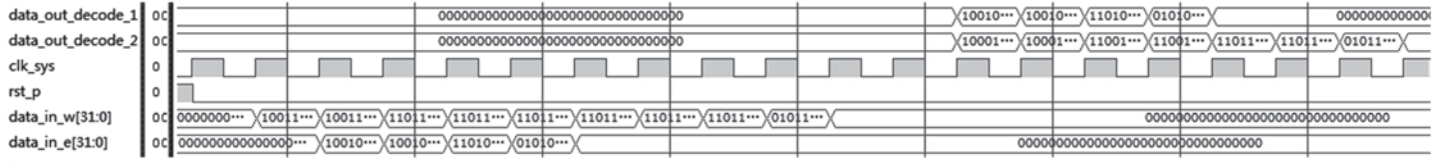
Figure 118.4a shows the comparison among unicast, multicast, and NC router. The data in the parenthesis for network coding router means the resource cost of

	Unicast Router	Multicast Router	Network coding Router
Used Registor	3675	3194	7214(1916)
Used LUTs	7043	6418	15312(8514)
Maximum Frequency(MHz)	162.515	193.611	184.298
Target Device	xc6vlx240t-1ff1156		

a



b



c

Fig. 118.4 **a** Hardware resource, **b** Simulation for coding operation, and **c** Simulation for decoding operation

the NC module. The new router costs double hardware resource and its frequency is moderate, as the NC module is added into the router. This is a slight compromise between costing and performance; however, with the rapid development of technique, the cost of hardware resource can be reduced.

Figure 118.4b and c is the simulation in terms of the coding and decoding operation. Figure 118.4b shows that two packets from the west and east direction are coded together to generate one packet. The total flit number of the two packets is 14. After the coding operation, the two packets turn to a coded packet, which contains only 10 flits. The number of flits in the coded packet is two (two new head flits) plus the number of flits in the longest packet in the coding operation. The blocking in the network is also mitigated to cut down the burden of the network. Figure 118.4c displays the decoding operation. The long packet is the coded packet, and the short one is the decoding packet. As the coded packet cuts off the head flit before the decoding operation, the decoded packet is two flits fewer than the coded packet.

118.5 Conclusion

This paper shows the design of a new NoC router based on NC, especially, the architecture of NC module is addressed. The router works in a distribution mode, which has enhanced the scalability of the NoC and simplified the design. The simulation result shows that NC module costs additional hardware resource overhead. With the research to be continued, new technique and theory will make the design simpler and less consumption of hardware resource.

Acknowledgments This work is supported by No. JCYJ20130401171935815 Shenzhen Research Funding, No. ZDSY20120619151314964 Shenzhen Innovative Environment Construction Scheme, and No. B08038 the 111 Project Grant. The authors would like to appreciate S'ebastien Le Beux for his insightful comments and suggestions, which have improved the quality of the paper. The authors also want to appreciate Gang Wu for many helpful discussions.

References

1. Jerger NE, Peh LS, Lipasti MH. Virtual circuit tree multicasting: a case for on-chip hardware multicast support, computer architecture, ISCA '08. 35th international symposium on, IEEE; 2008. p. 229–40.
2. Samman FA, Hollstein T, Glesner M. Adaptive and deadlock-free tree-based multicast routing for networks-on-chip, very large scale integration (VLSI) systems. IEEE Trans. 2010;18(7):1067–80.
3. Daneshtalab M, Ebrahimi M, Mohammadi S, Afzali-Kusha A. Low-distance path-based multicast routing algorithm for network-on-chips, computers & digital techniques. IET. 2009;3(5):430–442.
4. Ahlswede R, Cai N, Li S, Yeung R. Network information flow. IEEE Trans Inf Theory. 2000;46(4):1204–16.

5. Indrusiak LS. Evaluating the feasibility of network coding for NoCs, reconfigurable communication-centric systems-on-chip (ReCoSoC), IEEE. 2011. p. 1–5.
6. Duong-Ba T, Nguyen T, Chiang P. Network coding in multicore processors, international performance computing and communications conference (IPCCC), IEEE 30th International, IEEE. 2011. p. 1–7.
7. Shalaby A, Goulart V, Ragab ME-S. Study of application of network coding on NoCs for multicast communications, embedded multicore Socs (MCSoC), 2013 IEEE 7th international symposium on, IEEE. 2013. p. 37–42.
8. Vonbun M, Wallentowitz S, Feilen M, Stechele W, Herkersdorf A. Evaluation of hop count advantages of network-coded 2D-mesh NoCs, power and timing modeling, optimization and simulation (PATMOS), 2013 23rd international workshop on, IEEE. 2013. p. 134–41.
9. Hu J, Zhang S. NoC architecture with local bus design for network coding, communications and networking in China (CHINACOM), 2011 6th international ICST conference on, IEEE. 2011. p. 1151–4.

Chapter 119
An Improved Concurrent Programming Architectural Model Based on Cloud Computing

Wuxue Jiang, Yuqiang Chen and Minxia Liu

Abstract MapReduce, as one of the main concurrent programming models based on cloud computing, has become the research hot spot of information technology. Aiming at the development of MapReduce application program of high quality and efficiency, the working mechanism based on the Hadoop MapReduce model is analyzed in this chapter and MapReduce concurrent workflow is elaborated at the level of development class library, including task creation, job initialization, task initialization, communication between task and job. In addition, in order to solve the problem of Reduce input imbalance, a universal Map-Balance-Reduce improved model is proposed in this chapter. The balance layer embedded an adaptive splitting algorithm is added to the MapRduce model before reduce targeted at Reduce's defect of input imbalance, and its function is to guarantee the balanced Reduce input with the semanteme unchanged. The simulation indicates that the unbalanced degree of the improved MBR is obviously lower than that of MR; finally, some improvement prospects of the open source MapReduce model are discussed.

Keywords MapReduce model · Concurrent programming · HeartBeat program · Framework model · Map-balance-reudce

119.1 Introduction

Cloud computing is an inevitable development result of such technologies as distributed processing system, parallel processing, and grid computing for adapting to modern service needs. The cloud computing, as a new type of computing mode,

W. Jiang (✉) · Y. Chen
Department of Computer Engineering, Dongguan Polytechnic, 523808 Guangdong, China
e-mail: jiangwuxue_1@163.com

M. Liu
School of Electronic Engineering, Dongguan University of Technology,
523808 Guangdong, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_119

is also a new kind of computer resource composite mode and a kind of innovative business mode. As a new type of computing model and computer resource composite mode, it represents an innovative business model. MapReduce computing model draws upon the design idea of Lisp language applied in the field of artificial intelligence in early years, which adopts the idea of transparent details such as concurrent processing, fault tolerance, and data distribution; and designs the Map operation and Reduce operation, in which, the former is capable of processing each logic block in input record and generating a group of middle key value pair set while and the latter can merge data of key values with the same key [1,2]

The concurrent study on MapReduce computing model is based on the development of procedure-oriented language environment set, object-oriented multithreading set and message passing interface (MPI) environment set [3,4]; but their development principles are complex, and the development process is redundant and inefficient; thus in order to provide a concurrent computing of mass data processing of high quality and efficiency, Java class library set is analyzed for the implementation principles of Hadoop MapReduce and an universal prototype of MapReduce development framework is put forward to enhance the development efficiency of cloud computing program and the utilization ratio of cloud computing platform.

119.2 Mapreduce Model

MapReduce is a coarse-grained concurrent programming model and a new technology developed by Google company to process large database computing at PB order of magnitude on cheap computer cluster [5]. MapReduce model provides an interface of upper layer than multithreading programming, shields many detail problems of earlier concurrent computing and distributed processing in particular, reduces the entry requirements for developers, eases developers' burden, increases the development efficiency of this model, and the utilization ratio of distributed resources. MapReduce implementation process includes two principal aspects of mapping and protocol, and the information transfer there between is achieved by key value [6].

Hadoop MapReduce is a concurrent computing model operated in HDFS, which belongs to the client-server file system. One host node is named NameNode, and other slave nodes are named DataNode. MapReduce can decompose the job of NameData node JobTracker into multiple map tasks which are then mapped to the TaskTracker of each DataNode node. One map task can be performed in any computer node of cluster [7]. Multiple map tasks can be concurrently executed in multiple nodes in cluster. Map task is responsible for converting input record into key value. All the output of map task will be regrouped into the set of multiple orders, and a separate reduce task will be allocated to each set. Then the reduce task will process the key value of orders in set and multiple data values of associated key values. Reduce tasks are running concurrently in the cluster [8]. The implementation process of one MapReduce task and the type of data input and output can be ex-

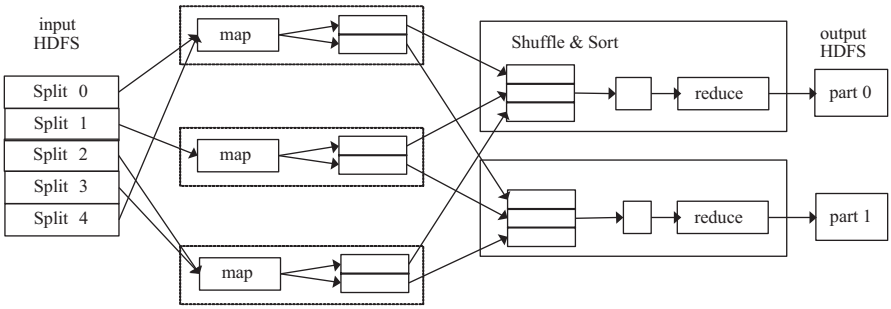


Fig. 119.1 MapReduce model diagram

pressed as follows: (input <k1,v1> → map → <k2,v2> → combine → <k2,v2> → reduce → <k3,v3>(output). MapReduce model is shown in Fig. 119.1.

119.3 Mapreduce of Class Library Development Level

The MapReduce's implementation of one job scheduling includes the following processes: create job, initialize job, and execute task with continuous communication process between task and job.

119.3.1 Job Creation

Firstly, one job example is created by JobConf class and then is configured. When JobClient.runJob (job) starts to run job and generate a JobClient example, the job is divided into several data subsets based on the specific InputForma class configured by JobConf, and each data set corresponds to one map task. The information such as path, offset, and block size of each data block is packed in JobFile.jar. JobClient submits tasks to JobTracker by its method submitJob(), and then JobTracker will follow relevant information in JobFile.jar to schedule and execute Job. Then the JobStatus object recording Job status information is returned to JobClient example. At meantime, JobClient creates a RunningJob object to regularly monitor and print the status information of JobStatus object, as well as monitor information.

119.3.2 Job Initialization

Job creation is followed by Job initialization. After JobTracker receives the Job, it will create one JobInProgress object which will manage and schedule tasks based on parameters such as Job.jar location and data to be processed by map task and

reduce task. All job objects as submitted are saved in one of member variable jobs in map, and the member variable jobsByPriority of list will apply the dynamic priority management to the submitted Job objects. JobTracker will schedule the highest priority job by JobInitThread() method, and start substantial job initialization after implementing initTasks() method.

119.3.3 Task Initialization

The JobInProgress object creates TaskInProgress object, by which the monitoring object of map and that of reduce will be created. Among them, the number of monitored objects by Map is decided by the number of segmentation blocks in JobClient. readSplitFile(), and that by ruduce is decided by the setting of JobConfig class. If not specified, the default will be an object monitored by reduce. Afterwards, JobInProgress will create JobStatus object to record information such as job implementation status and schedule. When JobTracker activates tasks by launchTask of each TaskInProgress, when Task object (i.e., Map Task and Reduce Task) is serialized and written in the service of TaskTracker, then the map task and ruduce task are started based on configuration.

119.3.4 Communication Between Task and Job

Upon the start of job and task and the implementation of divided labor, JobTracker is responsible for initialization and scheduling at job level, and TaskTracker for management and implementation at Task level. In the process, TaskTracker calls run() method recurrently to connect with JobTracker. And after the successful connection, TaskTracker calls heartBeat() method to build communication with JobTracker at intervals (the default is 10 s): firstly, the status information of executed task on this node by heartBeat() program is sent to JobTracker; and then under the impact of heartbeat program the response dispatched from JobTracker is obtained by transmitHeartBeat() method; finally, the instruction set transmitted from JobTracker is obtained by HeartbeatResponse.getActions() method. This set is stored in the form of array, and its content includes new task that JobTracker needs TaskTracker to perform and the tasks in progress for termination. Of course, before TaskTracker sends the heartbeat program to JobTracker, the number of tasks executed by this node and the utilization conditions of local disks should be monitored so as to judge whether the conditions for receiving new tasks are met, and then decide to set ask For New Task item of heartbeat program to be false or true based on conditions to control whether to accept the new task of JobTracker or not.

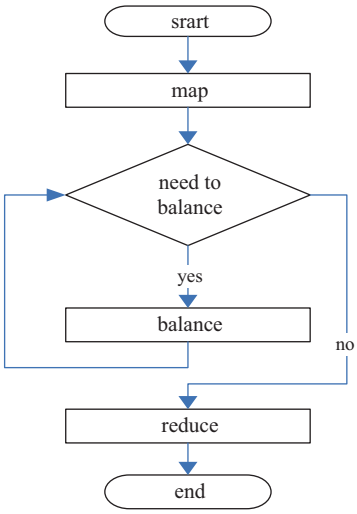


Fig. 119.2 Flow Diagram of MBR Model

119.4 Improved Mapreduce Model: MBR

119.4.1 Improved MBR Running Flow Diagram

One improved MapReduce model—Map-Balance-Reduce (MBR)—is put forward in this chapter. And the new model still remains simple, but the balance layer is added before reduce targeted at reduce's defect of input imbalance. The function of balance layer is to guarantee the balanced reduce input while keeping semanteme unchanged, so the users' required interfaces change from the two of map and balance into the three of map, balance, and reduce. The flow diagram of MBR architectural model is shown in Fig. 119.2.

119.4.2 The Improved MBR Program Framework Model

The balancer class is added to the improved MBR model framework, and its parameters are the same with those of Mapper and Reducer classes. The concrete model framework is shown as follows.

```
public class Mapper<KEYIN,  VALUEIN, KEYOUT, VALUEOUT>{
Protected void map(KEYIN key, VALUEIN value, Context context);}
public class Reducer<KEYIN, VALUEIN, KEYOUT,  VALUEOUT>{
protected void reduce(KEYIN key, Interable<VALUEIN> value, Context con-
text);}
public class Balancer<KEYIN, VALUEIN, KEYOUT,  VALUEOUT>{
Protected void balance(KEYIN key, VALUEIN value, Context context):}.
```

119.4.3 Balance Strategy

In this chapter, the balance strategy is an adaptive splitting algorithm. When the data blocks is written for every time, multiple DataNodes selected by the NameNode is used to store the current block. Whether the written splitting should be with the last splitting in the same shard is decided by the current record information written to the file. The concrete algorithm is described as follows:

```
Input: pFile, pp      Output: DNarr
1:  len=length[pfile.blocks]
2:  bi=len % pp
3:  if pFile!=null and bi>0 and bi>pp
4:     then block=pFile.blocks[len-1]
5:          DNarr=the DataNode list of block
6:  else    DNarr=chooseRadom()
7:  return  DNarr
```

119.5 Simulation Experiment Based on the CloudSIM

The Cloud computing simulation software—CloudSim is jointly released by the grid laboratory at University of Melbourne in Australia and Gridbus project group. CloudSim is extended on the original discrete event simulation package—SimJava has good cross-platform performance. In CloudSim platfrom, the method of DatacenterBroker class originally simply binds the task to the specified virtual machine running. In order to meet specific scheduling policy as required in this chapter, DatacenterBroker class is extended with its bindCloudletToVm() method overloaded; at the same time, the cloud task class—Cloudlet and virtual machine class—Vm are extended.

In this chapter, the MBR adopts comprehensive dynamic feedback load balancing scheduling algorithm [5], homework InputSize is 50M; the number of Maps, respectively,10, 25, 50, 100, and 200; the number of reduce is 1. The simulation test result is shown in Fig. 119.3.

It can be seen from the Fig. 119.3, the disequilibrium degree of M-B-R adopting the Balance MapRedue is lower than that of M-R. With the increase of the number of Maps, the advantage is more obvious.

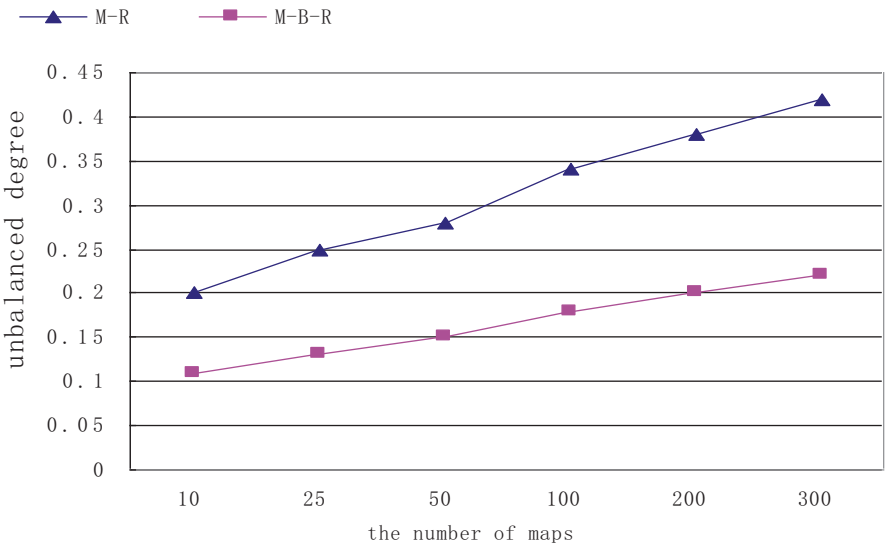


Fig. 119.3 Simulation result chart of unbalanced degree between M-R and M-B-R

119.6 Conclusion

The MapReduce programming model is one of core technologies of cloud computing software system, through which a simple business processing model is extracted from complicated implementation details; and a series of interfaces with powerful functions and easy accessibility are provided. MapReduce is not only a programming model but also an efficient task scheduling model. As a relatively popular cloud computing programming model, the application prospect of MapReduce is optimistic; however, some issues still exist such as simpler Job (task) scheduling algorithm, single message communication mechanism, and unclear computing and storage division. In order to improve the computing efficiency of MapReduce, hot issues of future MapReduce study will focus on the improvements in MapReduce task scheduling mechanism, the data storage strategies of communication with HDFS, the data input format and segmentation as well as the message communication monitoring.

Acknowledgments This work was financially supported by the National Natural Science Foundation of China (No.11047150), Guangdong Province University Outstanding Young Teachers' Training Program (No.Yq2013201), and by the Fund Project of Dongguan Polytechnic (No.2014a06, No.2014d16, No.2013c07, 2013b08).

References

1 Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters. Commun ACM. 2008;51(1):107–13.
2 Srirama SN, Jakovits P, Vainikko E. Adapting scientific computing problems to clouds using MapRedcue. Future Gen Comp Syst. 2012;28(1):184–92.
3 Barroso LA, Dean J, Hölzle. Web Search for a planet: the google cluster architecture. IEEE Micro. 2003;23(2):22–8.
4 Riedel E, Faloutsos C, Gibson GA, Nagle D. Active disks for large-scale data processing. Computer. 2001;134(6):68–74.
5 Srinivasa Rao P Reddy KT Prasad MHMK. A novel and efficient method for protecting internet usage from unauthorized access using map reduce. Int J Inf Technol Comp Sci. 2013;5(1):49–55.
6 Urbani J, Kotoulas S, Maassen J, Van Harmelen F, Bal H. WebPIE: a web-scale parallel inference engine using mapreduce. Web-Scale Semant Inf Process. 2012;10(10):59–75.
7 He Y, Tan H, Lu o W, Mao H, Ma D, Feng S, Fan J. MR-DBSCAN: an efficient parallel density-based clustering algorithm using mapreduce. In: Proc. IEEE International Conference on Parallel and Distributed Systems. New York, IEEE Press; 2011. p. 473–80.
8 Chen Y, Ganapathi A, Griffith R, Katz R. The case for evaluating mapreduce performance using workload suites. In: Proc. IEEE Annual International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems. New York, IEEE Press; 2011. p. 390–9.

Chapter 120
An Improved Character Recognition Algorithm Based on Shape

Lili Song and Jianfeng Han

Abstract Improvements have been proposed to overcome the shortage of existing algorithms based on the shape of the character. The background value assignment was adopted to describe the convex–concave characteristics, and the symmetry of convex–concave and the angle of concave region were proposed as supplement. The binary images of 25 uppercase characters and ten digits were identified correctly by the multilevel classifier. The new algorithm is immune to the size of character and tilt. Experimental results show that the improved algorithm effectively extends the range of identification, highlighting the advantages of shape-based algorithm.

Keywords Character identification · Shape · Convex–concave · Symmetry of convex–concave · Angle of concave region

120.1 Introduction

In the digital image, the target area is usually a form communicating with each other, a set of pixels with certain uniform consistency, and can be understood as a region. In general, the pixel value of a region is smooth and continuous. With the similarity criteria of these characteristics in the two-dimensional image, the pixels are grouped. Thus, the image plane is divided into a series of region containing the corresponding shape.

One of key technologies of character recognition is the feature extraction. In order to identify a character, we must describe the shape by extracting certain characteristics to represent the shape. The representation and description of shape can

L. Song (✉) · J. Han
School of Information Engineering, Inner Mongolia University of Technology, 010080 Huhhot, China
e-mail: songlili80@sohu.com

J. Han
e-mail: hanjianfeng@imut.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_120

be divided into two categories: the region based and the boundary based [1, 2], in which, the methods based on boundary, for example, describing the line, curve and boundary by chain code [3], are commonly used. These methods have great contour features of objects, but they are fragile to size, tilt and other factors of a character; consequently, it may need to be pretreated such as normalizing, tilt correction, and thinning. These processes are more tedious and will exert adverse effect on the identification of the exact character quickly. As to the other category, the methods based on region make full use of all the information from the edge to the shape of interior; therefore, these methods have a stronger ability of shaping characterization and robustness.

The feature of overall shape, such as convex–concave of the characters has been used to extract feature [4, 5]. Because of the immunity to size and tilt of character, this method omits the normalizing, tilt correction and thinning to improve the speed of character recognition; however, it is less complicated in extracting convex–concave feature to identify characters. So the complexity and storage space as required is relatively large. Such as calculation in transform domain [6], and computing histogram in logarithmic coordinates [7, 8] are needed. Other supplementary methods are demanded additionally, such as calculating the curvature to extract the feature of line and describing character by code to extract line and angular points.

As to this issue, new analysis and extraction algorithm based on convex–concave feature of characters is proposed in this chapter. And 25 uppercase characters (except "I") and 10 Arabic numerals (0–9) are recognized to show the feasibility and advantages of the approach.

120.2 Description of Convex–Concave

The convex–concave means the convex and concave characteristics of the character shape showed in the human eye. There are many ways to describe the convex–concave of characters and the background assignment method is used in this chapter. First, the binary image is scanned pixel by pixel. In addition, the background points (pixel value of 1) are selected as origin, from the origin as ray in the direction of the eight neighborhoods. Then the number of ray, which intersects with the foreground points (pixel values of 0) is assigned to the current background point. The pattern of Fig. 120.1a is the background assigned to the background assignment image as shown in Fig. 120.1b. Finally, the convex–concave of image is described by the new background value.

The background value of 8 represents all the eight rays from this point in eight directions intersect with foreground points, namely, the point is surrounded by strokes, the called loop area, as shown in Fig. 120.1a of the L1 and L2; the points background value for 5–7 are called concave area, as shown in Fig. 120.1a of the LC, RC, UC, DC. Obviously, the convex–concave characteristics of different characters have visible differences between others, and the robust is strong.

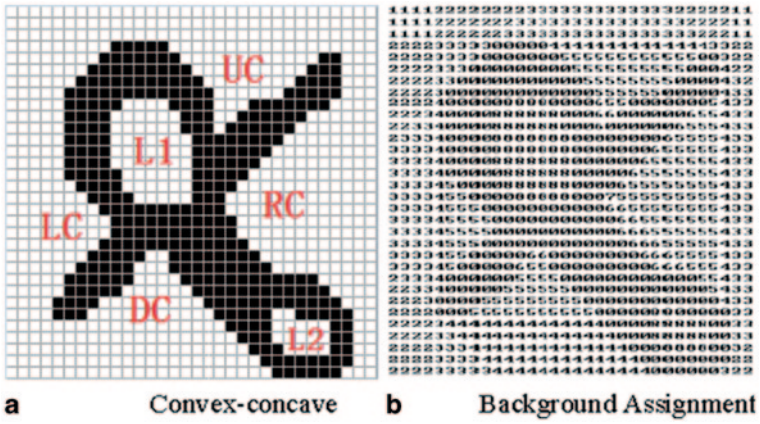


Fig. 120.1 Concave region and background assignment

In this chapter, seven characteristics, the number of loop regions; the number and direction of the convex–concave region; the size of the loop; the symmetry of convex–concave; the angle of convex–concave region; the location of convex–concave region and aspect ratio, are used for recognition.

120.3 Extraction of Convex–concave

120.3.1 Loop Region L

As mentioned above, the background value of loop region is assigned 8. But in fact, as the strokes encircle closely, there are deep concave region existing in some font characters; consequently, misjudgment of the loop may happen. The points on deep concave region are assigned 8 falsely according to simple background value assignment. For example, character "6" as shown in Fig. 120.2a. To avoid this mistake, the second-assignment is to additionally detect whether its eight neighbors in the loop for the points are assigned to 8. If the point is not in a loop, the value wound be reassigned to 9, and regard it as the concave region, as shown in Fig. 120.2b.

120.3.2 Concave Region C

Four different concave regions are shown in Fig. 120.1a. To distinguish them from each other, some definitions are shown as below:

Left-side concave LC: no foreground point on the left side;

Right-side concave RC: no foreground point on the right side;

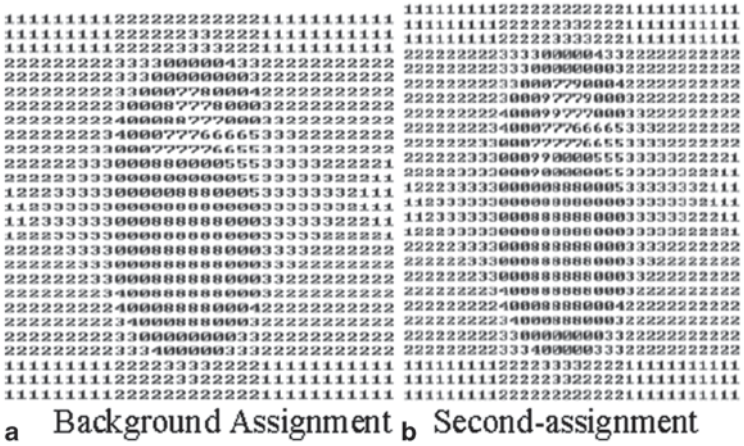


Fig. 120.2 Background assignment of "6"

Upside concave UC: no foreground point on the up side;
Downside concave DC: no foreground point on the down side.

120.3.3 Size of Loop SL

Many characters have loop, so the size of loop can be used to distinguish from different characters. For example, "D" and "O" have a bigger loop, and "4" and "R" have a smaller loop. SL could be defined as the ratio of the loop area to whole character area.

120.3.4 Symmetry of Convex–Concave CS

It is more complicated to identify the characters which have similar convex–concave characteristics, such as "2" and "Z", "5," and "S." Some papers adopt the calculation of the level intersections to distinguish them. This method has some limitations, such as the lower of "5" and "S" which have two intersections while the upper has one and two intersections; but it is too more difficult to accurately measure, and it is fragile to tilt. Interesting results have been found in studies that the upper and lower concave region of "S" and "Z" are symmetric in most common cases, while "5" and "2" are not symmetric. Therefore, CS is defined as the similarity of two concave regions of one character. For the sake of simplicity, CS is calculated by the ratio of the area of the upper and lower concave region of one character and CS = 1

indicates symmetry. CS can also be set by other value to represent the symmetric degree of convex–concave according to the specific circumstance.

120.3.5 Angle of Convex–Concave Region CA

Character "U" and "V", "H" and "N" are similar; but "V" and "N" have acute-angle concave region, i.e., an acute angle between the left and right strokes. "H" and "U" have the right-angle concave region, as shown in Fig. 120.3.

Suppose one pixel is in concave region. The pixel of its 8-neighborhood will be set 1 when it is a foreground point, and set 0 when it is a concave point. CA is defined as the sum of 8-neighborhood; so the pixel satisfied CA>5 must be found on acute angle concave region, and the pixel satisfied CA=5 must be found on the right angle concave region. None of them could be found on the concave region excluding the acute/right angle.

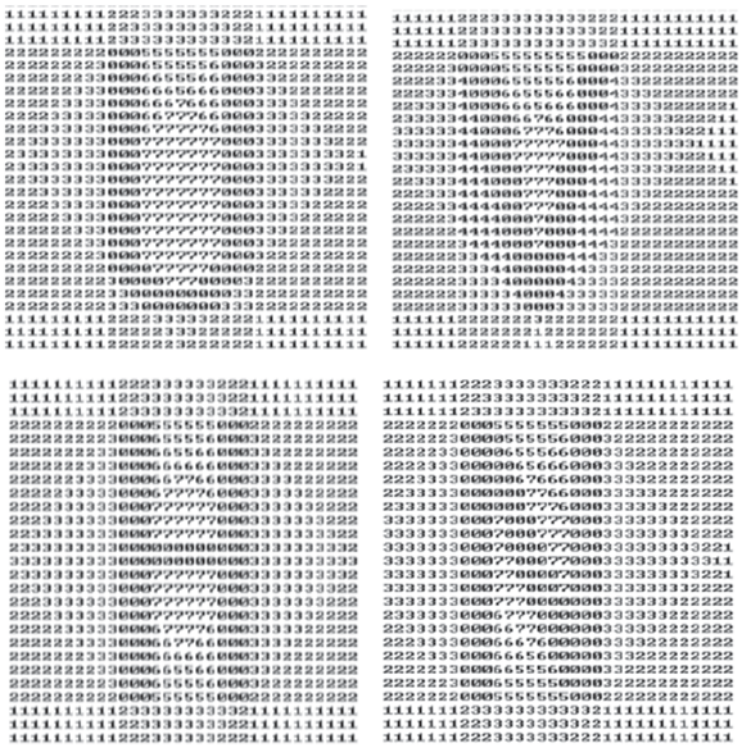


Fig. 120.3 Background assignment of "U," "V," "H," and "N"

120.3.6 Location of Convex–Concave Region CP

CP < 0.5 indicates concave region on the upper of character and CP > 0.5 indicates the lower.

120.3.7 Aspect Ratio CR

"0" and "O" could be differed by the aspect ratio. Suppose CR < 0.75 corresponds to "0," and CR > 0.75 corresponds to "O."

120.4 Experimental Results and Analysis

The multilevel classifier is designed to extract the convex–concave characteristics of 35 characters. One or more features are used by each level to classify the character in a step-by-step manner. Lots of experiments are performed with 32×32 binary images including the commonly used fonts in VC6.0 environment.

120.4.1 First-Level Classification

The characters are classified into three classes by the feature of no loop, 1 loop and 2 loops, as shown in Table 120.1.

120.4.2 Second-Level Classification

According to the convex–concave features of the first classification results, the three classes are classified into multi subclasses by multi conditions in the second-level classification, as shown in Table 120.2. As to Class 1, the subclass 1-1 is selected firstly by the left-side concave, and subclass 1-2 and 1-3 are selected by one and two right-side concaves, respectively. The subclass 1-4 is selected by the upside concave or the downside concave. Only "1" has no concave and belongs to subclass 1-5. Subclass 2-1 and 2-2 are classified by large or small loop. Class 3 only have two characters "8" and "B," they could be classified by the left-side concave obviously.

Table 120.1 First-level classification results

Class	Feature	Member
1	L=0	12357CEFGHJKLMNSTUVWXYZ
2	L=1	0469ADOPQR
3	L=2	8B

Table 120.2 Second-level classification results

Subclass	Feature	Member
1-1	LC≥1	2357JSTXYZ
1-2	RC=1	CGKL
1-3	RC=2	EF
1-4	UC≥1/DC≥1	HMNUVW
1-5	C=0	1
2-1	SL<0.7	469APR
2-2	SL>0.7	0DOQ
3-1	LC=1	8
3-2	LC=0	B

120.4.3 Third-Level Classification

The character unable to be distinguished in the first two levels needs the third-level classification, as shown in Table 120.3. In addition, “N,” “H,” “V,” “U,” “D,” “0,” and “O” are also classified by calculation of the angle of concave region.

120.4.4 Fourth-Level Classification

All characters are distinguished finally, as shown in Table 120.4.

Table 120.3 Third-level classification results

Sub-class	Feature	Member
1-1-1	RC=1,C=1,DC=1	X
1-1-2	UC=1	JY
1-1-3	LC=2	3
1-1-4	LC=1,CP<0.5,RC=1,CP>0.5	2Z
1-1-5	LC=1,CP>0.5,RC=1,CP<0.5	5S
1-1-6	RC=1	T
1-1-7	RC=0	7
1-2-1	UC=1,DC=1	K
1-2-2	UC=1	L
1-2-3	DC=1	G
1-2-4	UC=0,DC=0	C
1-3-1	DC=1	F
1-3-2	DC=0	E
1-4-1	UC=1,DC=2	M
1-4-2	UC=2,DC=1	W
1-4-3	UC=1,DC=1,CA>5	N
1-4-4	UC=1,DC=1,CA=5	H
1-4-5	UC=1,CA>5	V
1-4-6	UC=1,CA≤5	U
2-1-1	LC=1,RC=1	4
2-1-2	LC=1,RC=0	9
2-1-3	RC=1,DC=1	PR
2-1-4	LC=0,RC=1	6
2-1-5	LC=RC=UC=0,DC=1	A
2-2-1	RC=1/DC=1	Q
2-2-2	CA≥5	D
2-2-3	CA<5	0O

Table 120.4 Fourth-level classification results

Subclass	Feature	Member
1-1-2-1	CP<0.5	Y
1-1-2-2	CP>0.5	J
1-1-4-1	CS=1	Z
1-1-4-2	CS≠1	2
1-1-5-1	CS=1	S
1-1-5-2	CS≠1	5
2-1-3-1	C=1	P
2-1-3-2	C=2	R
2-2-3-1	CR<0.75	0
2-2-3-2	CR>0.75	O

120.5 Conclusion

The 35 characters are identified correctly by multilevel classification by means of the convex–concave characteristics of shape. The method is immune to the size of the characters and tilt. In particular, the use of symmetry of convex–concave and the angle of concave region effectively extends the range of identification. The proposed algorithm fully shows the superiority of the identification based on shape. As it can avoid complicated calculation and mathematical transformation, the improved algorithm could be much faster and convenient. It will be widely used in the recognition of printed and handwritten characters.

References

1. Loncarie S. A survey of shape analysis techniques. Pattern Recog. 1998;31(8):983–1001.
2. Zhang DS, Lu GJ. Review of shape representation and description techniques. Pattern Recog. 2004;37(1):1–19.
3. Freeman H. On the encoding of arbitrary geometric configuration. IEEE Trans Electric Comput. 1961;10(2):260–8.
4. Chen YK, Wang JF. Segmentation of single or multiple-touching handwritten numeral string using background and fore-ground analysis. IEEE Trans Pattern Anal Mach Intell. 2000;22(11):1304–17.
5. Kim CE. Digital convexity, straightness and convex polygons. IEEE Trans Pattern Anal Mach Intell. 1982;4(6):618–26.
6. Zhang DS, Lu GJ. Shape based image retrieval using generic Fourier Descriptor. Sign. Process Image Commun. 2002;17(10):825–48.
7. Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Trans Pattern Anal Mach Intell. 2002;24(4):509–22.
8. Zheng YF, Doennann D. Robust point matching for non-rigid shapes by preserving local neighborhood structures. IEEE Trans Pattern Anal Mach Intell. 2006;28(4):643–9.

Part X
Information Technology

Chapter 121
Fiber-Optic Extrinsic Fabry–Perot Interferometer Pressure Sensor Demodulation System with Three Quadrature Signals

Ningfang Song, Yujie Yang, Ying Chen and Jingming Song

Abstract A demodulation algorithm for fiber-optic extrinsic Fabry–Perot interferometer (EFPI) pressure sensor is proposed and demonstrated in this chapter. This method provides a wide range of pressure measurement and overcomes disadvantages of the traditional small sensing scale. The length of the sensing cavity could be recovered by using three quadrature phase-shifted signals. A theoretical description of the method is given here. An experimental sensing platform and calibrated behavior of the system were set up to prove the method. The results showed that the sensing accuracy is higher than 0.3 μm over a large range of 20 μm.

Keywords Extrinsic Fabry–Perot interferometer · Demodulation system · Pressure sensor · Quadrature phase-shifted signals · Large scope

121.1 Introduction

Optical fiber Fabry–Perot (F–P) sensors have been widely used due to their advantages of small size, light weight, and being insensitive to an electromagnetic interference. Extrinsic F–P cavity interferometer, as one structure of the sensor families, can be used for aerodynamic studies of complex pressure sensing. The sensors, along with appropriate algorithms, can be used to calculate a variety of critical parameters including stress, temperature, pressure, and operational load. Demodulation algorithms for extrinsic F–P interferometer (EFPI) sensors can be divided into two basic categories: intensity demodulation and phase demodulation [1–2]. The intensity demodulation method utilizes monochromatic light source. Physical perturbation is a function of the phenomenon being measured. It causes a change in received light intensity [3–4].

Y. Yang (✉) · N. Song · Y. Chen · J. Song
Institute of Optoelectronic Technology, Beijing University of Aeronautics and Astronautics, 100191 Beijing, China
e-mail: yyujie90@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_121

Most traditional methods for F–P sensor were useful in the measurement scope which is smaller than $\pm\lambda/8$ [5–6]. Because the interferometer transfer function of a fiber F–P cavity can be represented as a cosinusoidal function. That is, when the induced pressure changes its polarity, the F–P interferometers would not be able to detect this change if the switch took place at a maximum or minimum of the transfer function curve [7]. Our method is proposed by using three quadrature phase-shifted signals and continuous monitoring of the changes of phase difference. In this way, we will be able to resolve this issue and enlarge the measurement scope.

This chapter is organized as follows: after describing the principle of operation in Sect. 121.2, the experimental setup and results of the static pressure measurement are shown in Sect. 121.3. Finally, conclusion is drawn in Sect. 121.4.

121.2 Principle of Operation

A schematic of the EFPI sensor system with three quadrature signals unit is shown in Fig. 121.1. Our EFPI demodulation algorithm is based on an intensity demodulation method [8]. The system can get feedback from the pressure composed on F–P cavity by using three light sources.

The EFPI demodulation system should match the premise, while the three different transmission wavelengths must be properly selected to meet the quadrature phase-shifted relationship. The phase difference of these signals which have different wavelengths is an odd multiple of $\pi/2$ [9]. The $\pi/2$ differential phase shift in their outputs can provide quadrature signals. When these three chosen resonant wavelengths are $\lambda_1, \lambda_2, \lambda_3$, the phase at different wavelengths can be expressed as:

$$\phi_i = \frac{4\pi nL}{\lambda_i},\ i = 1,2,3, \tag{121.1}$$

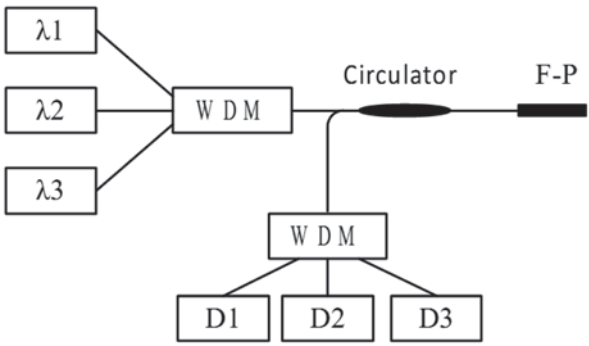


Fig. 121.1 Schematic of the EFPI sensor system. *λ1, λ2, λ3* represent the three light sources; *D1, D2, D3* are the optical power meters

Where n is the refractive index of the F–P cavity and its value is 1, because the cavity is filled with air; L is the length of the cavity; λ is the incident wavelength. The phase relationship between these three quadrature signals can be expressed as:

$$\phi_1 - \phi_2 = \phi_2 - \phi_3 = (2n+1)\frac{\pi}{2},\ n = 0,1,2\cdots. \tag{121.2}$$

For selecting of the light sources, the wavelength separation between the resonant wavelengths should be an odd multiple of $\frac{\lambda^2}{8nL}$ [10].

In our case of a low-finesse F–P sensor, the reflectivity of the interfaces R is about 4%, approximately, the reflected light intensity can be regarded as a two-beam interference [11]. In order to facilitate the calculation process, these three wavelengths should use the same intensity: $I_{10} = I_{20} = I_{30} = I_0$. The transfer function, in reflection of an EFPI cavity, is:

$$I_{iR} = I_0 \cdot 2R(1-\cos\phi_i), i = 1,2,3, \tag{121.3}$$

where ϕ is the light phase; I_0 is the incident light intensity.

Since $\phi_1 - \phi_2 = \phi_2 - \phi_3 = \frac{\pi}{2}$, the reflected light intensity can be rewritten as:

$$I_{1R} = I_0 \cdot 2R(1-\cos\phi_1) \tag{121.4a}$$

$$I_{2R} = I_0 \cdot 2R(1-\cos\phi_2) = I_0 \cdot 2R(1-\cos(\phi_1 - \frac{\pi}{2})) = I_0 \cdot 2R(1+\sin\phi_1) \tag{121.4b}$$

$$I_{3R} = I_0 \cdot 2R(1-\cos\phi_3) = I_0 \cdot 2R(1-\cos(\phi_2 - \frac{\pi}{2})) = I_0 \cdot 2R(1+\sin\phi_2). \tag{121.4c}$$

Using the trigonometric transformation formula, Eqs. (121.4a) and (121.4b) can recover the phase calculation formula for ϕ_1; similarly, using Eqs. (121.4b) and (121.4c) can recover ϕ_2. Through the reflected light intensity, the phase expressions can be expressed as follows:

$$\phi_1 = -\arctan(\frac{I_{2R}/I_0 - 2R}{I_{1R}/I_0 - 2R}) \pm m_1\pi,\ m_1 = 0,1,2\cdots, \tag{121.5a}$$

$$\phi_2 = -\arctan(\frac{I_{3R}/I_0 - 2R}{I_{2R}/I_0 - 2R}) \pm m_2\pi,\ m_2 = 0,1,2\cdots, \tag{121.5b}$$

where set the approximation $\lambda_1 \approx \lambda_2 \approx \lambda_3 \approx \lambda$. This approximation is valid within 1% of error for cavity lengths higher than 20 μm [12]. For a given cavity length, it

is always possible to define three resonant wavelengths that watch the above conditions. Giving the phase difference between λ_1 and λ_2 as ϕ_{12}, then

$$\phi_{12} = \phi_1 - \phi_2 = \frac{4\pi L}{\lambda_1} - \frac{4\pi L}{\lambda_2} = 4\pi L \cdot \frac{\Delta\lambda}{\lambda^2} \pm N\pi,\ N = 0\ or\ 1. \tag{121.6}$$

When the external pressure is applied, the cavity length will be changed. This will result in a variation in the phase difference.

$$\phi_{12}' = \phi_1' - \phi_2' = \frac{4\pi L'}{\lambda_1} - \frac{4\pi L'}{\lambda_2} = 4\pi L' \cdot \frac{\Delta\lambda}{\lambda^2} \pm N\pi,\ N = 0\ or\ 1. \tag{121.7}$$

In this way, the cavity length changes can be accurately obtained through the changes of the phase difference.

$$\Delta L = \Delta\phi_{12} \cdot \frac{\lambda^2}{4\pi\Delta\lambda} = \frac{\Delta\phi_{12} \cdot 2L}{\pi} \cdot \left(\Delta\phi_{12} = \phi_{12}' - \phi_{12}\right). \tag{121.8}$$

In order to demodulate the cavity length in a wide range, it is necessary to carry out a continuous monitoring of the changes of the phase difference between λ_1 and λ_2 [13]. Then, based on the signal of the hopping place, adding or subtracting $\neq$ can obtain the $\Delta\phi_{12}$. This method is proposed based on the traditional dual-wavelength method by adding a new incident light source and using new demodulation principle to solve the relative changes of the cavity length. In this way, the common mode error can be eliminated. The traditional dual-wavelength method by using two light sources solves the absolute cavity length; thus, the influence of the light source cannot be eliminated. Meanwhile, the traditional method must define the cavity length varied within $\pm\lambda / 8$.

Reflected light intensity at different wavelengths can be obtained according to the value of the cavity length L by theoretical calculating. Thus, using the above demodulation method can get two phases of the incident light source. By computing the changes of the two phases' deviation, varied cavity length can be obtained. Comparing the calculated cavity length with the given cavity length, the theoretical accuracy of the F–P sensor is given. Figure 121.2 shows the precision comparison chart between our demodulation method and the traditional dual-wavelength method [14]. Figure 121.2a shows our method's phase changes depending on cavity length and the precision calculation. Figure 121.2b is the traditional dual-wavelength method's precision calculation. In order to describe the measurement range and precision of the two methods more intuitively, the comparative data are illustrated by using a table (Table 121.1).

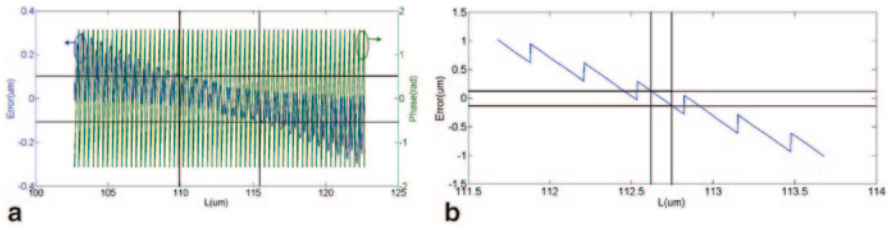


Fig. 121.2 Precision comparison chart between the three quadrature phase-shifted signals method and the traditional dual-wavelength method. **a** Precision calculation for three quadrature phase-shifted signals method. The *blue line* shows the theoretical accuracy; the *green line* represents the phase changes depending on the cavity length. **b** Precision calculation for traditional dual-wavelength method

Table 121.1 Precision calculation

	Measurement range	Precision	Measurement range	Precision
New method	5	0.1	20	0.3
Traditional method	0.3	0.1	Not applicable	Not applicable

The total length of cavity is 112.635 635 μm. The new method's cavity length varies more than 20 μm, and the theoretical accuracy is higher than 0.3 μm. When the F–P cavity length changes between 110 μm and 115 μm, the accuracy can reach to 0.1 μm. The traditional method's cavity length varies more than 2 μm, and the accuracy can reach 0.1 μm in the range of less than 0.3 μm ($\approx \pm \lambda/8$)

121.3 Experiments and Results

An experiment was designed to demonstrate the demodulation system. The schematic view of measurement system and extrinsic type of F–P are shown in Fig. 121.3. Three different wavelength lasers were used to illuminate an F–P cavity through a wavelength division multiplexing (WDM) and a circulator. The reflected interference signals were separated by using another WDM and detected by three optical power meters. Our EFPI sensor head was made of all silica technology that uses a commercial fusion splicer for fusing the single-mode incident and reflected fiber (single-mode fiber (SMF): 9/125 μm) into the capillary. All the fiber ends were angled cut to avoid spurious reflections.

In our experiment, the F–P cavity was installed on a hydraulic chamber. Small displacements were made in the cavity by changing the pressure of the hydraulic device. An SM125 (MOI Co. Ltd.) was used for static calibration, and its accuracy can reach 0.005 μm. The measurement system would realize the demodulation steps by using switch 1. The SM125 would calibrate the cavity length by using switch 2.

In the experiment, the original F–P cavity length is $L = 112.635$ μm. Three distributed feedback laser (DFB) diode lasers were used as sources. Different wavelengths were selected by using the formula: $\Delta\lambda = \frac{\lambda^2}{8nL}$, and the chosen resonant

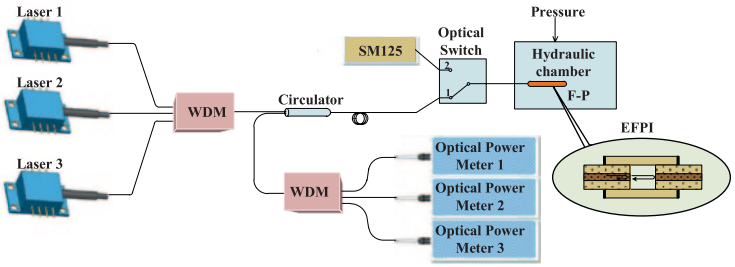


Fig. 121.3 Reflected light intensity measuring system

Table 121.2 Normalized reflected light intensity at different pressure

Pressure(map) I_R/I_o Wavelength(nm)	0	2	4	6	8	10
λ_1	0.0240	0.0593	0.0943	0.0322	0.0224	0.1553
λ_2	0.0230	0.1572	0.1587	0.0155	0.0240	0.1080
λ_3	0.1359	0.1014	0.0664	0.1265	0.1365	0.0056

wavelengths were 1308.096,1310.001 and 1311.904 nm. Values of normalized interferometer output intensity are shown in Table 121.2.

Figure 121.4a represents the phase changes over the cavity length. Figure 121.4b is the comparison chart between demodulation values through the measurement system, calibration results by using SM125, and the simulation results under the same point of cavity length, in which the black line represents the calibration results,

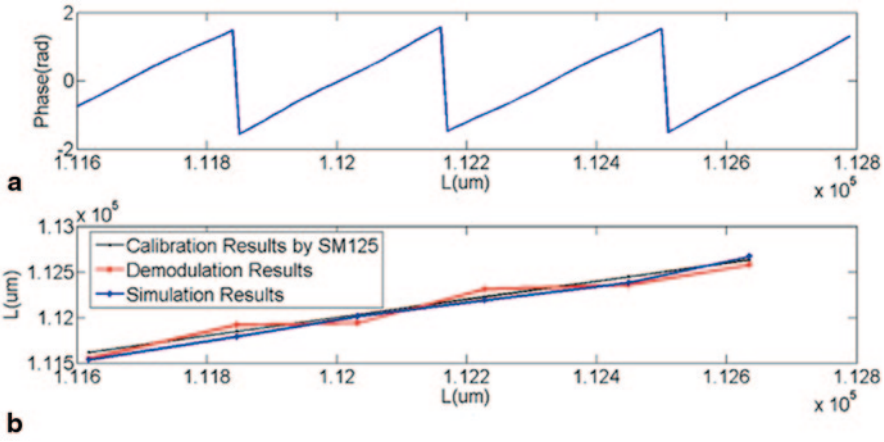


Fig. 121.4 Experiment results. The *black line* represents the calibration results by SM125, the *red line* represents the demodulation values through the measurement system, and the *blue line* represents the simulation results

the red line shows the demodulation values, and the blue line displays the simulation solutions. This figure shows the experimental results; it displays that a cavity length accuracy of 0.3 μm has been obtained. These data confirm the theoretically expected results. Because the two graphs are sharing the same horizontal axis, so we can see that this type of sensor system could be used for a wide range of pressure measurement. Although the three light sources are selected based on the initial cavity length, this method has no limit to the initial cavity length and according to the accuracy requirements to determine the range of the cavity. Because of the limited experiment conditions, the cavity length changes only at most 1 μm, but also exceeds $\pm\lambda / 8$.

121.4 Conclusion

In summary, a demodulation algorithm for fiber-optic EFPI pressure sensor is proposed. The premise is that the three different transmission wavelengths must be properly selected to meet the quadrature phase-shifted relationship. This method has successfully overcome the traditional dual-wavelength method's disadvantages of small sensing scale, providing a wide range of pressure measurement. The demodulation theory for F–P cavity length has been verified through simulation and experiment. The measuring accuracy can reach 0.3 μm in a large scope which is more than 20 μm/112.635 μm.

References

1. MacPherson WN, Kidd SR, Barton JS, Jones JDC. Phase demodulation in optical fibre Fabry-Perot sensors with inexact phase steps. IEE Proc-Optoelectron. 1997;144(3):130–3.
2. Zhang XM, Liu YX, Bae H, Pang C, Yu M. Phase modulation with micromachined resonant mirrors for low-coherence fiber-tip pressure sensors. Opt Express. 2009;17(26):23965–74.
3. Adachi M. Phase-shift algorithm for white-light interferometry insensitive to linear errors in phase shift. Opt Rev. 2008;15(3):148–55.
4. MacPherson WN, Gander MJ, Barton JS, Editors. Blast-pressure measurement with a high-bandwidth fibre optic pressure sensor. Meas Sci Technol. 1999;11(2000):95–102.
5. Wang JW, Lin RM, Guo DGT, Sun T. Development of a novel Fabry-Perot pressure microsensor. Sens Actuators A: Phys. 2004;116(2004):59–65.
6. Yu QX, Zhao XL. Pressure sensor based on the fiber-optic extrinsic Fabry-Perot interferometer. Photonic Sens. 2011;1(6):72–83.
7. Furstenau N, Schmidt M, Horack H, Goetze W, Schmidt W. Extrinsic Fabry-Perot interferometer vibration and acoustic sensor systems for airport ground traffic monitoring. IEE Proc Optoelectron. 1997;144(3):134–44.
8. Murpy KA, Gunther MF, Vengsarkar AM, Claus RO. Quadrature phase shifted extrinsic Fabry-Perot optical fiber sensors. Opt Lett. 1991;16(4):273–75.
9. Gander MJ, MacPherson WN, Editors. Embedded micromachined fiber-optic Fabry-Perot pressure sensors in aerodynamics applications. IEEE Sens J. 2003;3(1):102.

10. Lu E, Editor. Demodulation of micro fiber optic Fabry-Perot interferometer using subcarrier and dual-wavelength method. Optics Commun. 2011;285(7):1087–90.
11. Santos JL, Leite AP, Jackson DA. Optical fiber sensing with a low-finesse Fabry-Perot cavity. Appl Optics. 1992;31(34):7361–66.
12. Schmidt M, Furstenau N. Fiber optic extrinsic Fabry-Perot interferometer sensors with three-wavelength digital phase demodulation. Optics Lett. 1999;24(9):599–601.
13. Schmidt M, Werther B, Fürstenau N, Editors. Fiber-optic extrinsic Fabry-Perot interferometer strain sensor with <50 pm displacement resolution using three-wavelength digital phase demodulation. Optics Expr. 2001;8(8):475–80.
14. Dahlem M, Santos JL, Ferreira LA, Araújo FM. Passive interrogation of low-finesse Fabry-Pérot cavities using fiber Bragg gratings. IEEE Photon Technol Lett. 2001;13(9):990–92.

Chapter 122
Capacity Improvement by Multi-User Virtual Multi-Input and Multi-Output System in a Measured Indoor Environment at 5 GHz

Balla Moussa Coulibaly, Xiongwen Zhao, Suiyan Geng, Yiwei Li and Jingchun Li

Abstract In this chapter, wideband indoor multi-input and multi-output (MIMO) channel measurement was used to study the performance of the virtual MIMO formed by two users whose channels were measured while standing. In this measurement, 19 datasets are available which result in 19 users. From those datasets, we formed 171 possible virtual MIMO pairs. To set up a virtual MIMO system, two spaced antennas from standard users are taken and brought together. The capacities are evaluated for the standard users and the virtual MIMO system, and the comparison is made between standard users and the virtual MIMO formed by those users. The results show that the use of the virtual MIMO improves capacity in a great proportion compared to standard users. The parameters such as *K*-factor, root mean square (rms) delay spread, and spatial correlation, which affect the capacity are also evaluated. The results show that the capacity improvement is mainly due to spatial correlation and rms delay spread reduction rather than the *K*-factor reduction.

Keywords Capacity · *K*-factor · Rms delay spread · Spatial correlation · Virtual MIMO

122.1 Introduction

The multi-input and multi-output (MIMO) uses multiple antennas in mobile station (MS) side as well as in base station (BS) side [1]. To overcome spatial correlations that arise for using multiple antennas, the virtual MIMO can be used for that approach. In that case, a system can have fewer multipath components best for root

B. M. Coulibaly (✉) · X. Zhao · S. Geng · Y. Li
School of Electrical and Electronic Engineering, North China Electric Power University, 102206 Beijing, China
e-mail: huadian_zhaoxw@126.com

J. Li
State Radio Monitoring Center, 100037 Beijing, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_122

mean square (rms) delay-spread reduction [2], and also there will be much space between antennas from users, so the spatial correlation will be reduced and system capacity improvement will be yielded. Virtual MIMO has been used in some specific applications such as wireless sensor networks [3, 4].

In this chapter, we study the capacity improvement by multi-user virtual MIMO compared to the individual users that form the virtual MIMO. We used 19 standing measurements, from which cooperating pairs are chosen to set up the virtual MIMO.

Virtual MIMO has recently emerged as one of the most significant wireless techniques, as they can greatly improve the channel capacity and link reliability of wireless communications. Many papers have been dealing with virtual MIMO but in our knowledge they did not deal with wideband indoor cases. Contrary to them, in this chapter, indoor measurement datasets are used at a center frequency of 5.25 GHz with a 100-MHz signal bandwidth and BS antenna elements are set to 4.

The chapter is structured as follows. Section 122.2 shows the measurement system and campaigns. Section 122.3 shows the virtual MIMO system construction method and parameters. Section 122.4 presents the result, and Sect. 122.5 draws the conclusion.

122.2 Measurement System and Campaigns

The measurement campaign took place in the Nuclear Regulatory Commission (NRC) building in Ruoholahti, Helsinki, Finland. The environments covered the following propagation elements: wide indoor areas, open office environments, office rooms connected by a corridor, and meeting rooms as illustrated in Fig. 122.1.

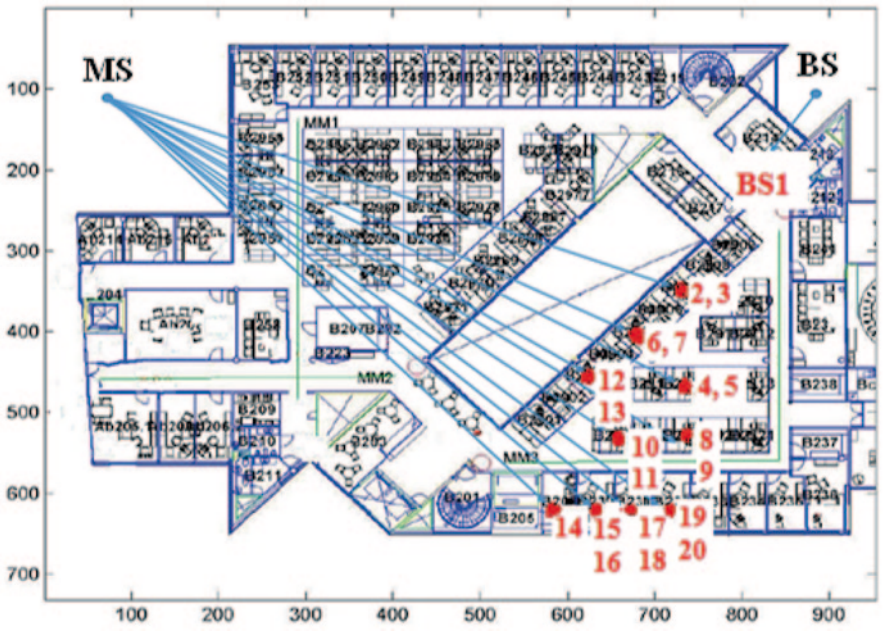


Fig. 122.1 Measurement locations in Finland

The measurements were performed using 5.25 GHz measurement frequency with 100 MHz radio frequency (RF) bandwidth. We used 19 standing datasets, from those datasets; a virtual 4 × 4 MIMO system was set up from two standard users of whom each activated only two of their antennas. The link considered here is the one from the BS to the two cooperating users.

122.3 Virtual MIMO System and Parameters

To set up a 4 × 4 virtual MIMO, two pairs of user will be selected to communicate with the BS. Two antennas from two users will be associated together to set up four Tx-antennas. As described above, the number of antennas in the BS is 24, but we only consider four antenna elements with four channels. The selection of antennas in each MS and BS must be done by taking advantage of low spatial correlation. So for that approach, we consider the two spaced antennas in each mobile user and the four spaced antennas in the BS. From mobile users, the two spaced antennas correspond to antenna 1 and 3 or antenna 2 and 4. In our calculation, we consider the two spaced antennas 2 and 4 for each user that correspond to the dashed column matrix 2 and 4 of the user channel matrices as shown in Fig. 122.2.Then the two dashed column matrix of the two users channel matrices are brought together to set up a virtual MIMO channel matrix. The properties of user 1 (U1) and user 2 (U2) will be analyzed and compared to the properties of the virtual MIMO formed by U1 and U2.

The key parameters studied in this chapter are capacity [5, 6], K-factor [7, 8], rms delay spread, and spatial correlation. The rms delay spread [9, 10] is evaluated using Eq. 122.1.

$$\tau_{rms} = \sqrt{\overline{\tau^2} - (\overline{\tau})^2} = \sqrt{\frac{\sum_i P_i \tau_i^2}{\sum_i P_i} - \left(\frac{\sum_i P_i \tau_i}{\sum_i P_i}\right)^2}, \tag{122.1}$$

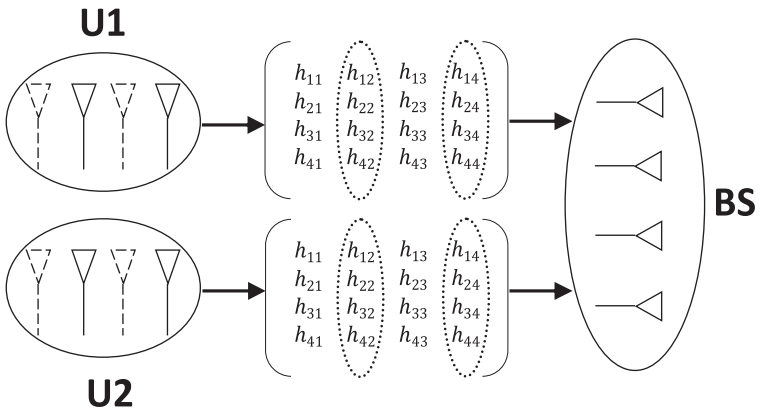


Fig. 122.2 Construction of 4 × 4 virtual MIMO system from standard users

where $\overline{\tau}$ is the mean excess delay, $\overline{\tau^2}$ is the mean of the square of excess delay, P_i is the received power (in linear scale) and τ_i is the corresponding excess delay of the i th path. And finally for a given channel matrix H, the correlation [11, 12] between the channels is defined as:

$$P_{ij,pq} = \frac{E_\tau\left\{h_{ij}(\tau)\,h_{pq}^{*}(\tau)\right\}}{\sqrt{E_\tau\left\{h_{ij}(\tau)\,h_{ij}^{*}(\tau)\right\}E_\tau\left\{h_{pq}(\tau)\,h_{pq}^{*}(\tau)\right\}}}, \tag{122.2}$$

where $E\tau$ is the expectation of over τ, and * denotes the complex conjugation. The average correlation coefficient of the channel, denoted by ρ, is computed by taking the mean value of all $\rho_{ij,pq}$.

122.4 Results

This section shows the results of all MIMO performance parameters such as the capacity, K-factor, rms delay spread, and spatial correlation.

122.4.1 Capacity

Figure 122.3a, b, and c show the comparison between the capacity of standard user 1, user 2, and the virtual MIMO. Three cases of capacities have been found using virtual MIMO: (a) virtual MIMO capacity is improved over both users (U1 and U2), (b) virtual MIMO capacity is improved over only one user (it can be either U1 or U2), and (c) virtual MIMO capacity degrades over both users (U1 and U2).

In the case where the virtual MIMO capacity is worse than both standard users, it has been found that at signal to noise ratio (SNR) of 30 dB, the losses are between 1 and 9 bit/s/Hz. Figure 122.3d shows that the capacity is much improved by a virtual MIMO over both standard users. There are some cases where the capacity is improved by a virtual MIMO over only one user. There are also only a few cases where the capacity is not improved by the virtual MIMO. Table 122.1 shows the range of data. It is found that the virtual MIMO achieves the maximum capacity, but the minimum capacity is also achieved in a small portion. Figure 122.4a shows the capacity of each individual standard user versus how many of possible virtual MIMO pairs can be offered to the users to improve the capacity of multi-user system. We have 19 measurement datasets which are equivalent to 19 users. Each user can have 18 partners. From each user, we will have 18 virtual MIMO pairs. The capacity of each user is compared to the capacity of its own 18 possible virtual MIMO pairs to see how many of the 18

possible virtual MIMO pairs would increase the capacity over each user. The capacity of the MIMO system is a dependent quantity. Capacity is affected by many

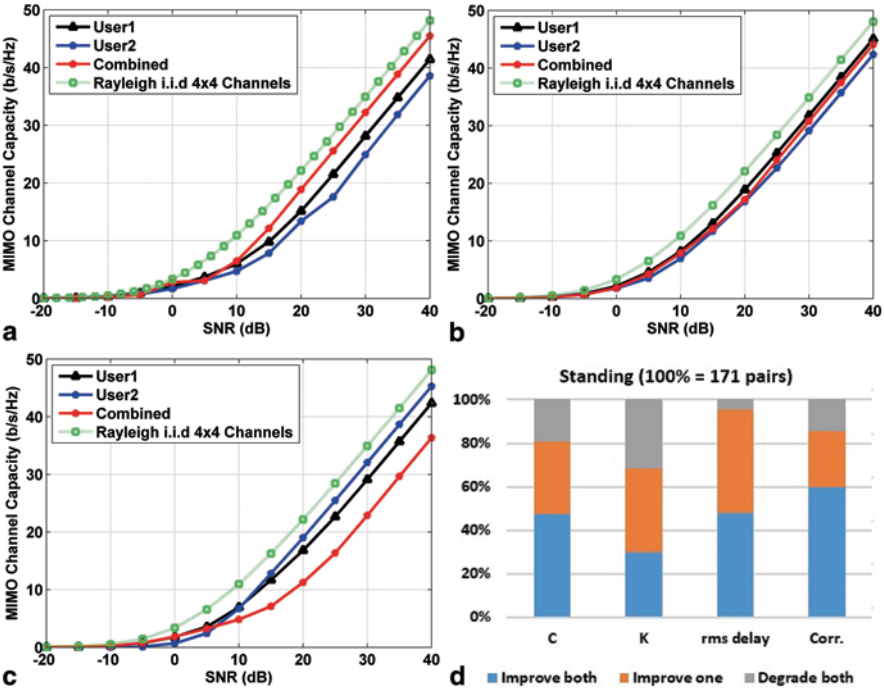


Fig. 122.3 **a** Capacity improved by the virtual MIMO over both users. **b** Capacity improved by the virtual MIMO over only one user. **c** Virtual MIMO capacity degrades over both users. **d** Percentage of capacity, *K*-factor, rms delay, and spatial correlation improving over both users or over only one user/degrading over both by the virtual MIMO system

Table 122.1 Range of capacity, *K*-factor and spatial correlation

Parameters	Classical MIMO	Virtual MIMO
Capacity, bps/Hz	18.42–32.050	17.81–36.11
K-factor	−3.29–11.29	−3.82– 3.95
Rms delay	77.64–81.74	77.44–81.42
Correlation	0.387–0.480	0.347–0.51

other channel parameters such as *K*-factor, rms delay spread, correlation, etc. These parameters will be the key study in what follows and will be analyzed in a similar way as capacity.

122.4.2 K-factor

From Table 122.1, the standard MIMO system has maximum *K*-factor + 11.29 dB, and we noticed that the virtual MIMO system lowered the *K*-factor in a great proportion.

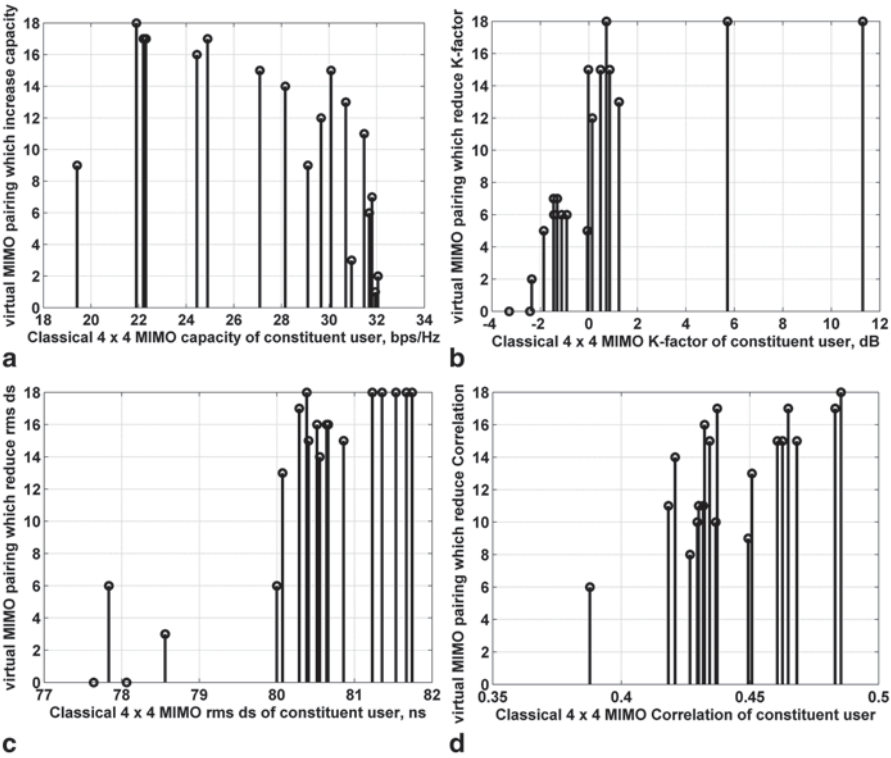


Fig. 122.4 **a** Capacities of users versus how many of the 18 possible virtual pairings could increase capacity for all pairs. **b** *K*-factor of users versus how many of 18 possible virtual channels pairs could reduce *K*-factor for all pairs. **c** rms delay spread of users versus how many of 18 possible virtual channels pairs could reduce rms delay spread for all pairs. **d** Correlation of users versus how many of 18 possible virtual channels pairs could reduce correlation for all pairs

The virtual MIMO system tries to make the K-factor positive where the K-factor of the standard users is negative. Figure 122.3d also shows the proportion reduction of the K-factor over both standard users and over one standard user. There are some cases where the K-factor is

not reduced by the virtual MIMO system but these cases are small compared to the cases where the K-factor is reduced over both standard users and over one standard user. Figure 122.4b, similar to Fig. 122.4a, shows the number of possible virtual pairings which can reduce the K-factor compared to an individual standard user. From Fig. 122.3d, the improvement of K-factor over both is about 30%. The percentage of the capacity improvement by virtual MIMO is much considerable compared to the percentage reduction of the K-factor by the virtual MIMO. So, it is difficult to mention that the capacity improvement is only due to the K-factor reduction. However, the contribution of rms delay spread and spatial correlation reduction by the virtual MIMO can be considered.

122.4.3 RMS Delay Spread

From Table 122.1, the maximum value of rms delay spread is seen with standard users. Of course, the virtual MIMO falls below the two standard users as depicted in Fig. 122.3d from which the virtual MIMO system reduces the rms delay spread in great quantity below both standard users as well as only one standard user. Only in a few cases, the virtual MIMO system increases the rms delay spread over both standard users. However, the improvement of virtual MIMO capacity is greatly due to the rms delay spread reduction rather than K-factor reduction. Figure 122.4c shows the rms delay spread of standard users versus how many of 19 possible virtual MIMO channels pairs could reduce the rms delay spread for all pairs. If we consider one standard user, there exist more or equal to three possible pairs among 18 pairs which results in the reduction of the rms delay spread. The exception is seen with the standard user 1 and user 3. So, it is difficult to say that the improvement in capacity is due only to the K-factor and rms delay spread. The contribution of the correlation reduction by virtual MIMO can be considered.

122.4.4 Spatial Correlation

From Table 122.1, in some cases, the virtual MIMO reaches the maximum value. But the effect of this is not that much because the correlation is increased only by about 0.05. In a similar way to the K-factor, Fig. 122.4d shows the number of possible virtual pairing which can reduce correlation compared to an individual standard user. There are more or equal to six possibilities that virtual MIMO pairing can reduce correlation compared to its specific standard user. As shown on Fig. 122.3d, in many cases the correlation is reduced by the virtual MIMO. The percentage of reduction in the correlation by the virtual MIMO is much higher than the percentage of reduction in the K-factor and the rms delay spread. The capacity improvement by the virtual MIMO is mainly due to the effect of correlation reduction rather than that of the K-factor and rms delay spread reduction. Reducing correlation in a MIMO system has long been known to increase system capacity [13, 14].

122.5 Conclusion

The results in this chapter are derived by using a real multi-user MIMO measurement dataset, which shows the performance parameters of the virtual MIMO system compared to the standard user. It is found that the virtual MIMO system can provide capacity improvement. The results show that the system capacity is improved by about 47.3 % over both standard users and 33 % over only one standard user. But there are few cases where the virtual MIMO capacity degrades over both constituent users, which means that the capacity is not always improved by the virtual MIMO

system. However, the virtual MIMO is considered to be desirable because it tends to lift the capacity where the capacity is lowered by the standard user system. Also, considering one standard user, there exist more or equal one pairs that can improve its capacity. The *K*-factor is improved in a few cases by the virtual MIMO over both standard users, but it is improved much over only one standard user. However, contrary to the *K*-factor, the rms delay spread gives 48% improvement over both standard users by the virtual MIMO system, and the spatial correlation is improved by about 59.6% by the virtual MIMO over both standard users. It is found that only in few cases the rms delay spread and the correlation are increased by the virtual MIMO system. The results show that the improvement of capacity in the virtual MIMO system is more related to spatial correlation and rms delay spread reduction than the *K*-factor reduction.

References

1. Rusek F, Persson D, Lau BK, Larsson EG, Marzetta TL, Edfors O, Tufvesson F. Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process Magazine. 2013;30(1):40–60.
2. Payami S, Tufvesson F. Delay spread properties in a measured massive MIMO system at 2.6 GHz. 24th International Symposium on Personal indoor and Mobile Radio communications(PIMRC) London United Kingdom; 2013. p. 53–7.
3. Jayaweera SK. V-BLAST-based virtual MIMO for distributed wireless sensor networks. IEEE Trans Commun. 2007;55:1867–72.
4. Xu K, Chizuni D. A V-BLAST based virtual MIMO transmission scheme for sensor network lifetime maximization. IEEE Vehi Technol Conf. (Fall, Baltimore, MD); 2007. p. 377–81.
5. Brown T, Carvalho ED, Kyritsy P. Capacity of MIMO channel, in the MIMO Radio Channel with Matlab Examples, John Wiley & Sons, Ltd., Publication; 2012. p. 17–73.
6. Cho YS, Kim J, Yang WY, Kang CG. MIMO: channel capacity, in MIMO-OFDM Wireless Communications with Matlab, John Wiley & Sons Pte Ltd, Singapore; 2010. p. 263–80.
7. Ilic-Delibasic M, Pejanovic-Djurisic M. Impact of Random K factor on Ricean Fading Wireless System Performance, Electrotechnical Conference New York; 2012. p. 233–36.
8. Greenstein LJ, Ghassemzadeh SS, Erceg V, Michelson DG. Ricean K-factors in narrow-band fixed wireless channels: theory, experiments, and statistical models, IEEE Trans Vehi Technol. 2009;58:4000–12.
9. Rappaport TS. Wireless communications—principle and practice, 2nd edn. Prentice Hall, Upper Saddle River; 2002.
10. Zhao X, Kivinen J, Skog K, Vainikainen P. Characterization of Doppler spectra for mobile communication at 5.3 GHz. IEEE Trans Vehi Technol. 2003;52:14–23.
11. Hui HT, Wang X. Building antenna characteristics into MIMO channel simulation. Int J Electr. 2010;97(6):703–14.
12. Hai L, Zhang Y. Calculating correlation coefficients in 3-D mix-diversity MIMO channel with Kronecker product. Nanjing China, WCSP 2009; 2009. p. 1–6.
13. Shiu DS, Foschini GJ, Gans MJ, Kahn JM. Fading correlation and its effect on the capacity of multielementantenna systems. IEEE Trans Commun. 2000;48:502–13.
14. Chuah CN, Tse DNC, Kahn JM, Valenzuela RA. Capacity scaling in MIMO wireless systems under correlated fading. IEEE Trans Inf Theory. 2002;48:637–50.

Chapter 123
Selection of Test Points Based on Heuristic Graph Search

Xiaofeng Lv, Wen Ye, Xinghai Wang and Wei Pang

Abstract As to the problem of test point selection for the weapon system, the chapter firstly studies the correlation model and the information entropy theory to transfer the test point selection problem into the graph searching problem; then, the evaluation function is defined as the heuristic information, which deduces the graph expansion way and guides the map search; finally, the method is verified by experiment. Experiment results show that the method can improve the searching efficiency of test nodes and provide an effective guidance for optimal test point selection of the weapon system.

Keywords Test point selection · Correlation matrix · Information entropy · Heuristic graph search

123.1 Introduction

With the performance and complexity of the weapon system, the costs of maintenance and security have become increasingly high; thus, the development of the testability design of system is a task which brooks no delay. The premise of testability design is the test point selection; therefore, selecting appropriate test points to obtain an equipment state has become the focus of current research. Reference makes the group number fuzzy by defining a test point to get an optimal test set using the implicit method [1]. Reference designs the evaluation standard for measuring the test points so as to solve the optimal test set by using the implicit method

X. Lv (✉) · W. Ye · X. Wang · W. Pang
Naval Aeronautical and Astronautical University, 264001 Yantai, China
e-mail: lxf_00112@163.com

W. Ye
e-mail: 19804473@qq.com

X. Wang
e-mail: xinhai@sina.com

W. Pang
e-mail: 719970894@qq.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_123

[2–4]. Reference uses the boolean logic for the test point selection problem by combining the exclusion method and the implicit method, thus excluding redundant test set [5]. Reference optimizes the correlation matrix model by using the genetic algorithm and then uses the particle swarm optimization algorithm to search the optimal test point set [6]. Reference works out a minimum test set by using the method of decomposing matrix step by step and analyzing the weighted value of test points [7].

The above methods are adopted to find the optimal test point set under the condition of assuming the same test cost in the test points; but it is almost impossible that it happens in an actual system; therefore, this chapter presents a test point optimization method based on the heuristic graph search by firstly defining the information entropy of test points and evaluation function of test cost, then defining the extension rule of graph nodes to search out the optimum test point set at best test cost, and finally verifying this searching algorithm by examples.

123.2 Preliminaries

1. Correlation mathematical model

Correlation refers to the logic relationship existing between the component unit and the test point of unit under test (UUT), two component units, or two test points. The correlation mathematical model of UUT is described by the correlation matrix, as shown in Formula (123.1):

$$D_{m\times n}=\begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & & \vdots \\ d_{m1} & d_{m2} & \cdots & d_{mn} \end{bmatrix}, \tag{123.1}$$

in which the matrix of row i is:

$$F_i=[d_{i1}\ d_{i2}\ \cdots\ d_{in}]. \tag{123.2}$$

It represents the response information of the ith component unit fault's reaction in each test point information. While the matrix of column j is:

$$T_j=[d_{1j}\ d_{2j}\ \cdots\ d_{mj}]^T, \tag{123.3}$$

It represents the components' fault information of each test point, which can be measured. d_{ij} is the boolean variable, namely, if the fault F_i can be detected by testing point T_j, then $d_{ij}=1$, otherwise $d_{ij}=0$.

2. Information entropy of test points

Entropy represents an uncertain degree of system, and information measure is the entropy change before and after the information receipt [4]. Many references evalu-

ate the information content of test points by using entropy [2–7]. This chapter uses information entropy to evaluate the information of graph node by combining the heuristic graph searching method.

According to the information theory, assume X, Y to be discrete random variables, then entropy of X is:

$$H(X) = -\sum_{i=1}^{m} P(X_i) \log_2^{P(X_i)}. \tag{123.4}$$

The conditional entropy of X towards Y is:

$$H(X|Y) = -\sum_{i=0}^{m} \sum_{j=0}^{n} P(Y_j) P(X_i|Y_j) \log_2^{P(X_i|Y_j)}. \tag{123.5}$$

Information entropy about X provided by Y, namely the mutual information entropy between X and Y is:

$$I(X,Y) = H(X) - H(X|Y). \tag{123.6}$$

In the correlation matrix, assume $F = \{f_0, f_1, \cdots, f_m\}$ is the fault set of system, P is the probability of fault isolation, then the entropy of F is:

$$H(F) = -\sum_{i=0}^{m} p(f_i) \log_2^{p(f_i)}. \tag{123.7}$$

The available test set of system is $T = \{t_0, t_1, \cdots, t_n\}$, if the test point t_j divides the system into k subsets, encoding $1, 2, \cdots, k$, respectively. According to the formula of total probability, the conditional entropy of F can be expressed as:

$$\begin{aligned} H(F \mid t_j) = {} & p(F_{i0}) H(F \mid F_{i0}) + p(F_{i1}) H(F \mid F_{i1}) + \\ & \cdots \quad + p(F_{ik}) H(F \mid F_{ik}). \end{aligned} \tag{123.8}$$

Therefore, the information entropy about F provided by t_j is:

$$I(F, t_j) = H(F) - H(F|t_j). \tag{123.9}$$

In order to calculate conveniently, this chapter deducts another expression. As

$$p(F_{ir}) = \sum_{f_i \in F_{ir}} p(f_i), \tag{123.10}$$

so

$$H(F|F_{ir}) = -\sum_{i=0}^{m} p(f_i|F_{ir}) \log_2^{p(f_i|F_{ir})}. \tag{123.11}$$

In the formula $(0 \le r \le k)$ and because

$$H(F|T_j) = -\sum_{i=1}^{m} p(f_i)\log_2^{p(f_i)} + \sum_{f_i \in F_{j1}} p(f_i)\log_2^{p(F_{j1})} + \sum_{f_i \in F_{j2}} p(f_i)\log_2^{p(F_{j2})} + \cdots + \sum_{f_i \in F_{jk}} p(f_i)\log_2^{p(F_{jk})}, \tag{123.12}$$

so

$$I(F,t_j) = H(F) - H(F \mid t_j) = -\{p(F_{j1})\log_2^{p(F_{j1})} + p(F_{j2})\log_2^{p(F_{j2})} + \cdots + p(F_{jk})\log_2^{p(F_{jk})}\}. \tag{123.13}$$

Considering the testing cost, $p(F_{jr})$ is defined as the ratio of least test cost of isolating subset Min_r and total test cost of the left test points, namely $p(F_{jr}) = \dfrac{Min_r}{C_r}$. $I(F,t_j)$ represents the information entropy of the test points. The smaller the information entropy is, the less the testing cost for fault isolation will be.

123.3 Heuristic Graph Search

123.3.1 Graph Search About Test Points

Heuristic graph search is to sort extension table nodes by using heuristic information of the problem, thus to achieve the aim of reducing the search range and reducing the complexity of the problem [9, 10]. The test point selection process explains the process of graph nodes.

The extension rule of graph nodes is defined as follows:

1. To define the alternative test point set S_t and add all test points to alternative test point sets. Expand the available test point set of t_i which is $N_t = \{t \mid t \in (S_t - t_j)\}$; t_j is the expanded test point.
2. Assume the maximum fault number which is $N_{t_i \max}$ in the fuzzy set for the expanded test point t_i, then the least test points of fuzzy isolation is $\log_2^{N_{t_i \max}}$. If $\log_2^{N_{t_i \max}} > Num\{N_t\}$ and $Num\{N_t\}$ represents the number of elements in the set $\{N_t\}$, then t_i will not be extended.

123.3.2 Heuristic Evaluation Function

The test point selection is a problem of combinatorial optimization, which can be described by using the set covering model; but the set covering is N-P complete problem. When the system scale is large, it is difficult to achieve the optimal

solution; therefore, the heuristic evaluation function is introduced in this chapter in order to speed up the test point search process.

The heuristic evaluation function of test point t is defined as follows:

$$f(t) = g(t) + h(t). \tag{123.14}$$

In this formula, $g(t)$ is the sum of test cost of the path from node t to root node s; $h(t)$ is the least test cost from the root node to the target node.

Define the function:

$$f^*(t) = g^*(t) + h^*(t). \tag{123.15}$$

In this formula, $g^*(t)$ is the testing cost of node t to root node s, $h^*(t)$ is the test cost from the root node to the target node. Apparently, when the test node t extends, there is $g(t) = g^*(t)$ and $h^*(t) \geq h(t)$; clearly $f^*(t) \geq f(t)$, then using $f(t)$ as graph searching algorithm of heuristic evaluation function belongs to the A^* algorithm [11]. The test point selection problem under the condition of least cost in this chapter belongs to the limited map. The path exists from the initial node s to the target node t. The algorithm A^* can definitely find an optimal path to reach the target node [12]. The optimum test set under best test cost is the test point set from the root node to the target node.

123.3.3 Test Point Selection Based on Heuristic Search Algorithm

The A^* algorithm is to expand nodes through inspiring information provided by means of the evaluation function $f(t)$, then finding the best route to the destination node. This chapter uses the A^* algorithm in terms of the test point selection problem with the specific steps shown as follows:

Step 1: Initiate the graph node s and add all the test points to the collection S_t.

Step 2: Determine whether the fault is completely isolated and whether the evaluation value is a minimum; if yes, then all test root nodes tracing pointer from node t_x root node on the path is the best test point set. Then end the search, otherwise, go to Step 3.

Step 3: Get test points of minimum $f(t)$ value from the collection S_t; if there are several test nodes, then select the test point with minimum entropy as the expanding node and delete the node from set S_t. To check whether the fault is completely isolated and whether the evaluation value is the least; if yes, go to Step 2, otherwise, go to Step 4.

Step 4: Calculate node values of collection again according to Formula (123.14) and go to Step 2.

123.4 Simulation and Verification

This chapter uses the SK15HE gas turbine unit of conference [8] as an example to state the test point selection method of heuristic search graph. The correlation matrix is shown in Table 123.1, in which $f_1 \sim f_6$ represent different faults, $t_1 \sim t_7$ are available test points, and c is the test cost.

The first layer: Initialize the node s, add all test points to the alternative test point set S_t, and calculate heuristic evaluation value $f(t)$ for each test point value.

For test point t_1, fuzzy maximum fault number $N_{t_1 \max} = 4$, minimum test points $M_{t_1} = \left\lceil \log_2^4 \right\rceil = 2 < Num\{"N_t"\}$, so $f(t_1) = 5$. Delete t_1 from set N_t, then N_t turns to $\{t_2, t_3, t_6, t_4, t_5, t_7\}$.

For test point t_2, fuzzy maximum fault number $N_{t_2 \max} = 3$, minimum test points $M_{t_2} = \left\lceil \log_2^3 \right\rceil = 2 < Num\{"N_t"\}$, so $f(t_2) = 4$. Delete t_2 from set N_t, then N_t turns to $\{t_3, t_6, t_4, t_5, t_7\}$.

For test point t_3, fuzzy maximum fault number $N_{t_3 \max} = 3$, minimum test points $M_{t_3} = \left\lceil \log_2^3 \right\rceil = 2 < Num\{"N_t"\}$, so $f(t_3) = 5$. Delete t_3 from set N_t, then N_t turns to $\{t_4, t_5, t_6, t_7\}$.

For test point t_4, fuzzy maximum fault number $N_{t_4 \max} = 4$, minimum test points $M_{t_4} = \left\lceil \log_2^4 \right\rceil = 2 < Num\{"N_t"\}$, so $f(t_4) = 5$. Delete t_4 from set N_t, then N_t turns to $\{t_5, t_6, t_7\}$.

For test point t_5, fuzzy maximum fault number $N_{t_5 \max} = 5$, minimum test points $M_{t_5} = \left\lceil \log_2^5 \right\rceil = 2 > Num\{"N_t"\}$, so do not expand t_5.

For test point t_6, fuzzy maximum fault number $N_{t_6 \max} = 3$, minimum test points $M_{t_6} = \left\lceil \log_2^3 \right\rceil = 2 < Num\{"N_t"\}$, so $f(t_6) = 7$. Delete t_6 from set N_t, then N_t turns to $\{t_5, t_7\}$.

For test point t_7, fuzzy maximum fault number $N_{t_7 \max} = 4$, minimum test points $M_{t_7} = \left\lceil \log_2^4 \right\rceil = 2 > Num\{"open"\}$, so not to expand t_7.

Extended root node is shown in Fig. 123.1.

Second layer: Obviously, the estimated value of t_2 is $f(t_2) = 4$, which is the minimum value. Put $t_1, t_3, t_4, t_5, t_6, t_7$ into set "S_t", then set "N_t" is $\{t_1, t_3, t_4, t_5, t_6, t_7\}$, and continue to expand in this way.

Table 123.1 Correlation matrix of SK15HE gas turbine

	t_1	t_2	t_3	t_4	t_5	t_6	t_7
c	3	1	2	2	2	4	1
f_1	1	1	1	1	0	1	0
f_2	0	1	0	1	0	1	0
f_3	1	0	0	0	0	0	0
f_4	0	0	0	1	0	0	1
f_5	0	0	1	1	1	1	0
f_6	0	1	1	0	0	0	1

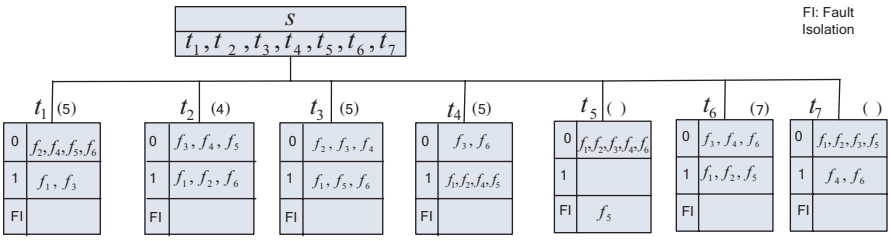


Fig. 123.1 Root node graph

Third layer: As $f(t_3)=f(t_4)=4$, calculate $I(F,t_3)=w_{t_3}\log_2^{w_{t_3}}$, in which $w_{t_3}=\dfrac{c_{t_7}}{c_{t_1}+c_{t_4}+c_{t_5}+c_{t_6}+c_{t_7}}$, thus $I(F,t_3)=\dfrac{1}{12}\log_2^{\frac{1}{12}}=0.2897$. $I(F,t_4)=w_{t_4}\log_2^{w_{t_4}}$, $w_{t_4}=\dfrac{c_{t_3}}{c_{t_1}+c_{t_3}+c_{t_5}+c_{t_6}+c_{t_7}}$, thus $I(F,t_3)=\dfrac{2}{12}\log_2^{\frac{2}{12}}=0.4308$; expand t_3 with the expansion method as shown above. The final complete extended search graph is shown in Fig. 123.2.

As can be seen in Fig. 123.2, the optimal test point set is $\{t_2,t_3,t_7\}$, and the test cost is 4.

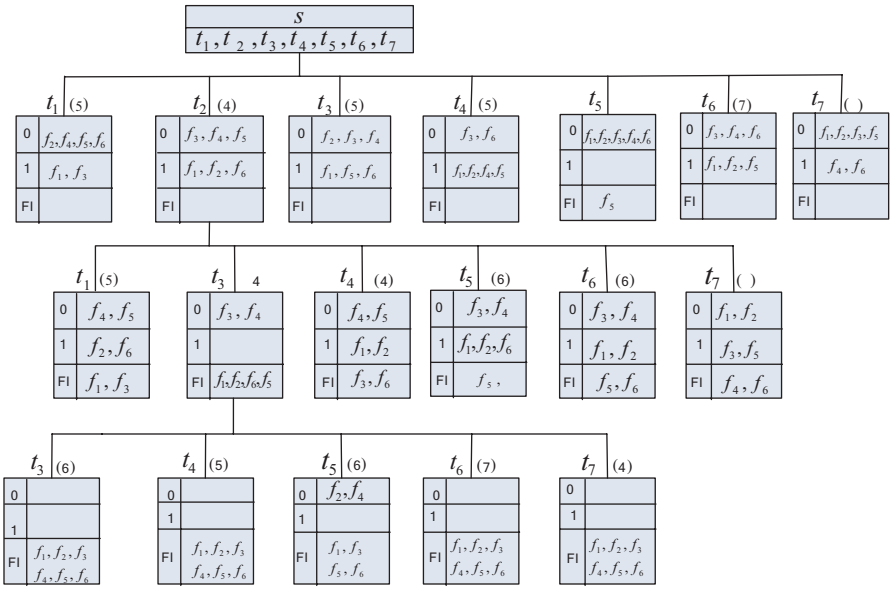


Fig. 123.2 Complete expansion graph

123.5 Conclusion

In this chapter, the problem of test point selection based on correlation matrix of electronic system is discussed, and the test point selection problem is transferred into the searching graph node problem. To avoid the blindness of graph node search, the heuristic strategy is adopted based on the heuristic evaluation function, which is derived from the correlation model and the information entropy theory; at meanwhile, the cost of each testing point is considered when all fault isolation rate (FIR) satisfies the system requirement. Experiment results show that this method can manage to search optimal test point sets quickly and effectively while satisfying the requirement of system testability. The proposed method has practical guidance significance for fault diagnosis as well as complicated electronic systems.

References

1. Prasad VC, Rao Pinjala SN. Fast algorithms for selection of test nodes of an analog circuit using a generalized fault dictionary approach. Circuit Syst Signal Process. 1995;14(6):707–24.
2. Prasad VC, Babu NSC. Selection of test nodes of for analog fault diagnosis in dictionary approach. IEEE Trans Instrum Meas. 2000;49(6):1289–97.
3. Pinjala KK, Bruce CK. An approach for selection of test points for analog fault diagnosis. Proceedings of the 18th IEEE International Symposium on defect and fualt tolerance in VLSI systems, IEEE; 2003. p. 287–94.
4. Starzyk JA, Liu D, Liu Z-H, et al. Entropy-based optimum test nodes selection for analog fault dictionary techniques. IEEE Trans Instrum Meas. 2004;53(3):754–61.
5. Yang P, Qiu J, Liu G. The test selection algorithms based on boolean logic. J Test Meas Technol. 2007;21(5):386–90.
6. Chen X, Qiu J, Liu G. Optimal test selection based on hybrid BPSO and GA. Chinese J Sci Instrum. 2009;30(8):1674–79.
7. Prasad VC, Babu NSC. Selection of test nodes for analog fault diagnosis in dictionary approach. IEEE Trans. Instrum Meas. 2000;49(4):1289–97.
8. Zhong T, Junyou S. System testability design analysis and demonstration. Beijing: Beijing University of Aeronautics and Astronautics Press; 2003. p. 48–54.
9. Diestel R. Graph theory. New York: Springer-Verlag; 2000. p. 89–96.
10. Yang C, Tian S, Long B. Application of heuristic graph search to test points selection for analog fault dictionary techniques. IEEE Trans. Instrum Meas. 2009;58(7):2145–58.
11. Pattipati KR, Alexandridis MG. Application of heuristic search and information theory to sequential fault diagnosis. IEEE Trans Systems Man Cybern. 1990;20(4):872–87.
12. Xiao W. Graph theory and algorithm. Beijing: Aviation Industry Press; 1993. p. 57–66.

Chapter 124
A Novel Method of Frequency-Selective Surface Transmission Characteristics Test

Jianwei Lian, Liang Chen, Haiyan Yan and Yanlong Li

Abstract A novel method of frequency-selective surface (FSS) transmission characteristics test is presented on the basis of the absorbing materials window of metal backboard which is derived from the characteristics test of finite size FSSs in engineering technology. The design method is validated by the numerical simulation (the finite-element method) as well as the experimental results.

Keywords Frequency-selective surface · Absorbing materials window of metal backboard · Transmission characteristics

124.1 Introduction

In the past few decades, the frequency-selective surfaces (FSSs) are developed and investigated by many researchers for their spatial filtering characteristics, and they have been widely applied in a variety of applications. Several typical examples of these applications include radomes, dichroic reflectors, polarizers, and absorbers, etc. [1–3]. Although FSSs have been attracted much attention, most of the calculations and simulations of FSS are focused on infinite FSSs, while finite size FSSs are applied in engineering; therefore, it is highly valuable to design a novel method of finite size FSSs transmission characteristics test which has high accuracy and precision.

In this chapter, a novel method is proposed to solve the problem. It is obtained by the simulations of different FSS test settings and validated by the simulation and experimental results. The following sections describe the simulation and validation of the test method.

J. Lian (✉) · L. Chen · H. Yan · Y. Li
Science and Technology on Electromagnetic Compatibility Laboratory, China Ship Development and Design Center, 430064 Wuhan, China
e-mail: Lianjianwei89@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_124

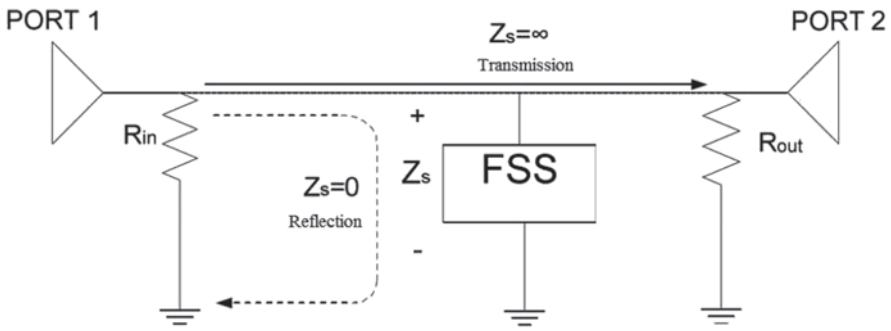


Fig. 124.1 Equivalent two-port network model of frequency-selective surface (*FSS*)

124.2 Test Theory of Frequency-Selective Surface

The test theory of FSS can usually be explained with the two-port network model, from the perspective of circuit [4]. As FSS is periodic structure, it can be modeled and analyzed with only one unit. Figure 124.1 has shown the equivalent two-port network model of FSS.

In Fig. 124.1, port 1 and port 2 represent the plane wave incident area and transmission area, which are separated by FSS. Rin and Rout are the input and output impedance, which are embodied by the air characteristic impedance.

As the input and output impedance of port 1 and port 2 are constant, the lower the surface impedance Z_s of FSS is, the more energy will be reflected back to port 1; the higher the surface impedance Z_s becomes, the more energy is transmitted to port 1.

124.3 Simulation and Analysis of Test Settings

Figure 124.2 shows the measurement system for obtaining the transmission characteristics of finite size FSSs. The system is formed by the signal source, the spectrum analyzer, the receive and transmit antennas, FSS test window, and the absorbing materials, in which the settings of FSS test window are particularly important in achieving the desired measurement precision; therefore, in the following section, two methods with different FSS test windows are researched to design the perfect settings of the FSS test window.

It is necessary to note that the main subject of this chapter is to simulate the test method of finite size FSSs transmission characteristics rather than the simulation of infinite FSSs; in this sense, the simulation conditions should be almost exactly the same as the conditions shown in Fig. 124.2. The free space is modeled with radiation boundary in high-frequency-selective surface (HFSS); the transmit and receive antennas are modeled with horn antenna, and the performance of the absorbing materials in HFSS is the same as the experimental conditions.

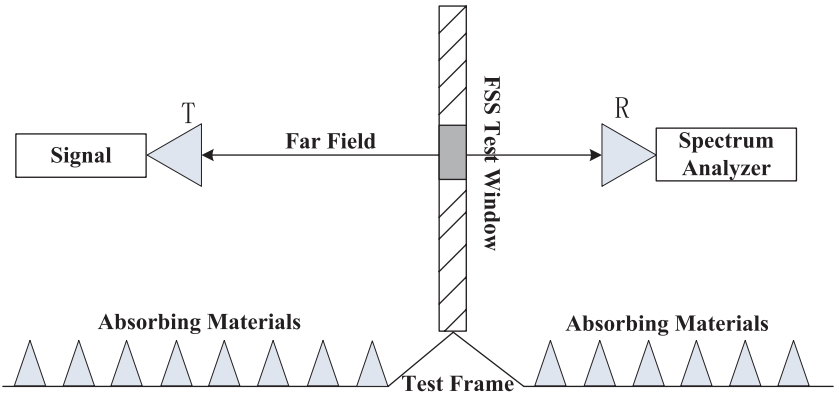


Fig. 124.2 Schematic diagram of frequency-selective surface (*FSS*) transmission characteristics test

Besides, the experimental size of FSS model is not completely the same with the simulation model due to the limitation of the computer's performance. But the result is still convincing because the transmission characteristics of the simulated model are stable when the finite size FSSs consist of 5×5 units, not less than this.

124.3.1 Modeling and Simulation for the Submodule of the Test Method

The submodules of the measurement system shown in Fig. 124.2, including infinite FSSs, horn antenna, and absorbing materials, are designed and simulated through HFSS before the simulation of two different test methods.

Figure 124.3 presents the simulated parameters for each submodule, where the infinite FSSs are composed of the ring-aperture unit, and its design resonant frequency is 1 GHz; the voltage standing wave ratio (VSWR) of horn [5–6] antenna is less than 1.7, its gain is about 12 dB, and the reflection of absorbing materials [7] is less than −30 dB. The performance follows the guidance of the experimental modules.

124.3.2 Modeling and Simulation of Test Method

Two different test methods, type-I and type-II, are studied in this section. Both of them are obtained by arranging two antennas and the models of finite size FSSs in between, as is shown in Fig. 124.4 and Fig. 124.6. The distance between each antenna and models meets the far-field requirements with the only difference between these two methods represented in the settings of the FSS test window.

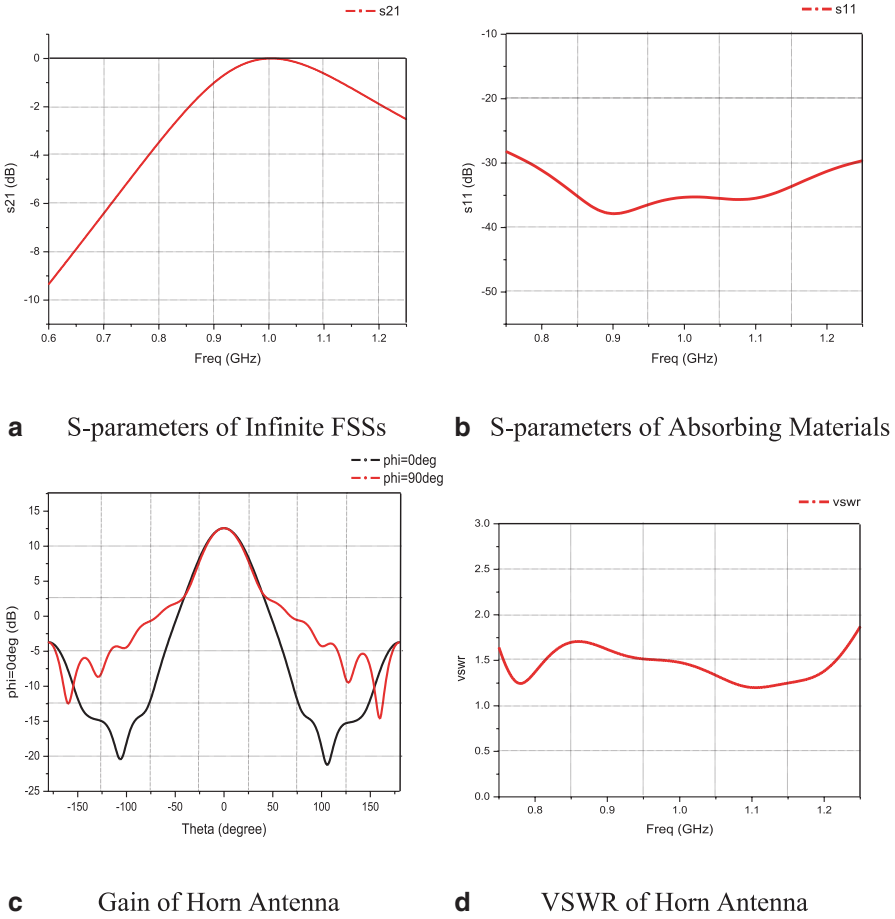


Fig. 124.3 Performance of each submodule. **a** S-parameters of infinite FSSs. **b** S-parameters of absorbing materials. **c** Gain of horn antenna. **d** VSWR of horn antenna

1. Type-I method: no FSS test window

As it can be seen in Fig. 124.5a, the simulation results with type-I method are considerably different from the infinite FSSs, especially for 0.95 to ~1 GHz, where the value exceeds zero. It is completely impossible for FSS. A closer inspection of the

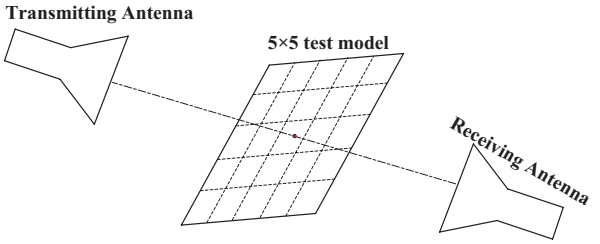


Fig. 124.4 Geometry graph of type-I method

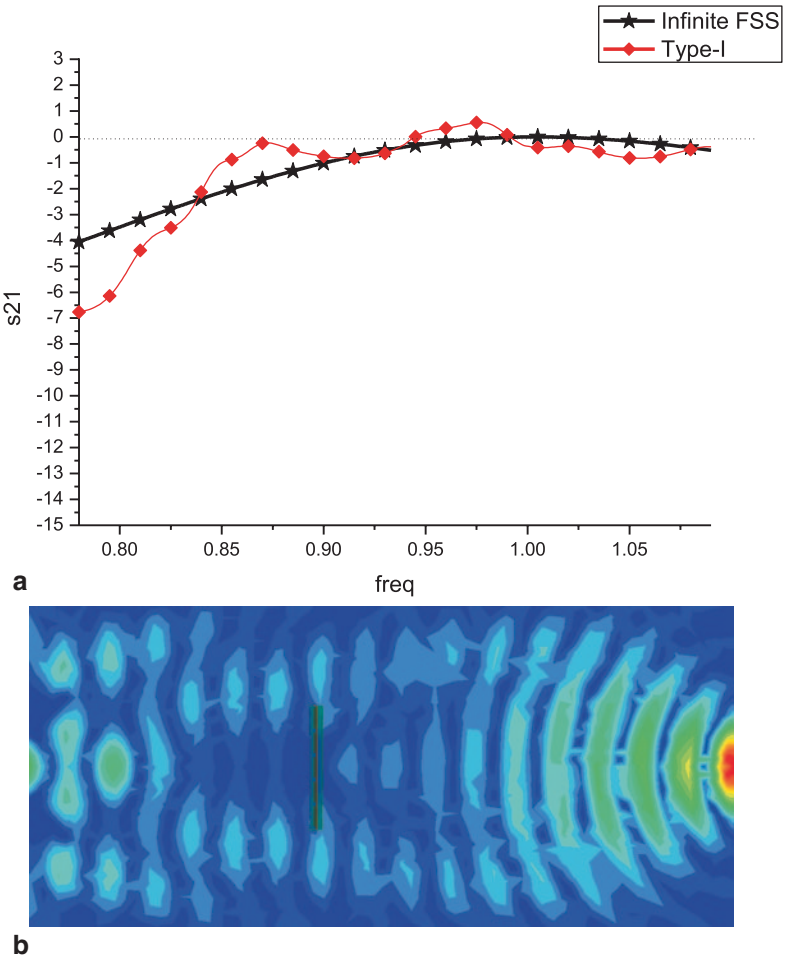


Fig. 124.5 **a** Finite element method (FEM) simulation results using type-I method, infinite frequency-selective surfaces (*FSSs*) simulation is also shown for comparison. **b** Electric field distribution

electric field distribution in Fig. 124.5b suggests that the presence of the edge diffraction is, in fact, responsible for that incorrect result.

2. Type-II method: the absorbing materials window of metal backboard

To effectively suppress the multipath effects and block the field edge diffraction, the absorbing material window of metal backboard is used here to compensate for the limited role of absorbing materials.

Figure 124.7 presents the HFSS simulation diagram of type-II method, where the boundary is set to radiation boundary to model free space and the transmission coefficients can be obtained from the S-parameters of the antenna.

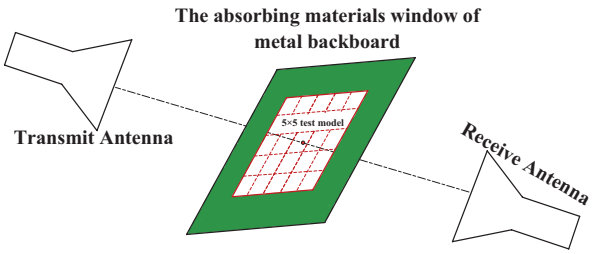


Fig. 124.6 Geometry graph of type-II method

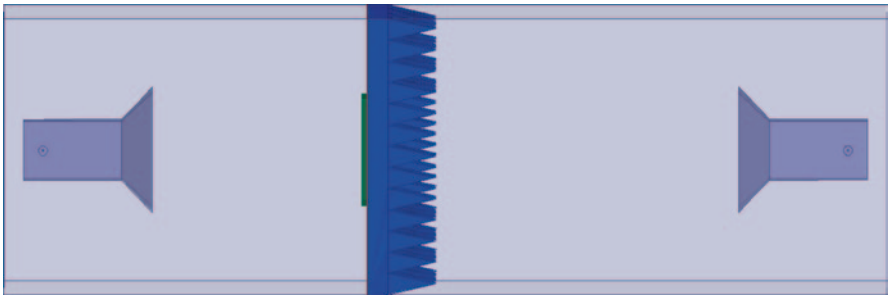


Fig. 124.7 Simulation schematic diagram of frequency-selective surfaces (*FSSs*) with type-II method

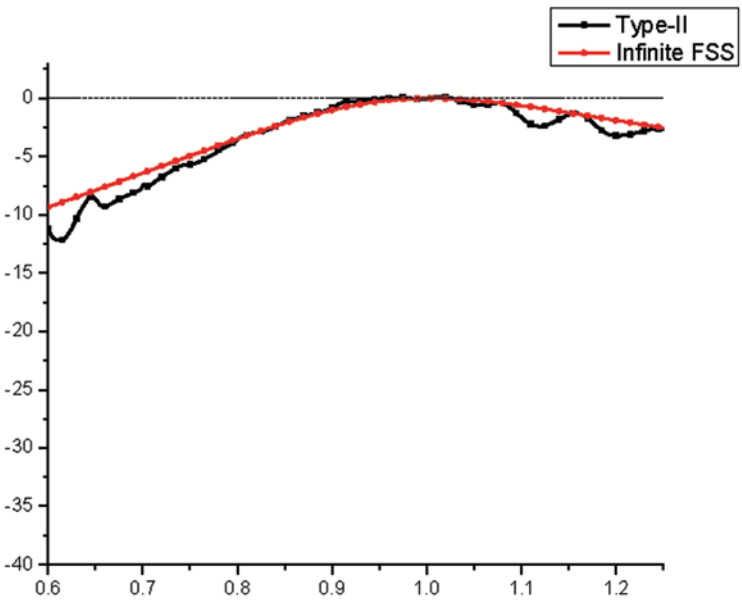


Fig. 124.8 Finite element method (FEM) simulation results using type-II method, infinite frequency-selective surfaces (*FSSs*) simulation is also shown for comparison

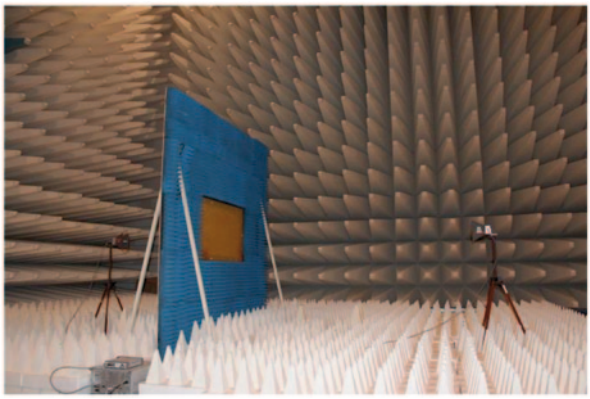


Fig. 124.9 Test system of type-II method

With the type-II method, the error can almost be neglected, which can be seen in Fig. 124.8 because the edge diffraction has been blocked with this method.

124.4 Validation of the Test Method

To ensure the measurement precision, the test has been carried out in the microwave anechoic chamber [8], where the ground is covered with the absorbing materials in order to reduce the multipath effects. The background is below −35 dBsm.

Figure 124.9 is about the test system of type-II method. The model under test is placed in the middle of these two antennas.

The measured and simulated transmission coefficients of type-II method are demonstrated in Fig. 124.10. The measured data are given for 0.6 to ~1.25 GHz band, which is the range of operation of the measurement setup. The measurement shows a perfect match with the HFSS simulation for 0.8 to ~1.1 GHz band; therefore, the type-II method can be adopted to achieve the transmission characteristics of finite size FSSs.

124.5 Conclusion

The absorbing materials window of metal backboard has been proposed as a novel method to obtain the transmission characteristics of finite size FSSs with high accuracy. The effectiveness of this method stems from the fact that the field edge diffraction can be blocked by type-II method; besides, the measured and simulated frequency response has been used to prove its validation. This novel method is expected to find great application values.

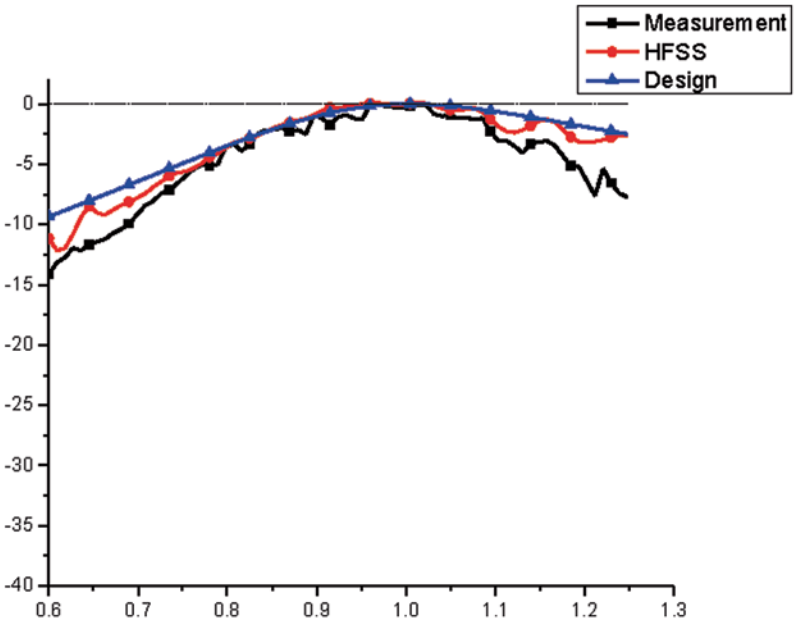


Fig. 124.10 Measured and simulated frequency response of type-II method, infinite FSSs are also shown for comparison

References

1. Gao Q, Yan DB, Yuan NC. Properties analysis of loaded frequency selective surface. Chinese J Radio Sci. 2006;21(3):441–4 (In Chinese).
2. Courtney CC. Time-domain measurement of the electromagnetic properties of materials. IEEE Trans MTT. 1998;46(5):517–2.
3. Friedsam GL, Biebl EM. A broadband free-space dielectric properties measurement system at millimeter wavelengths. IEEE Trans. 1997;46(2):515–8.
4. Munk BA. Frequency selective surfaces. New York: Wiley; 2000. p. 284–293.
5. Shao YF. A simplified design of wide b and ridged horn antenna. Mod Radar. 2004;26(3):65–7 (In Chinese).
6. Bruns C, leuchtmann P, Vahldieck R. Analysis and simulation of a 1-18 GHz broadband double-ridged horn antenna. IEEE Trans Electromagn Compat. 2003;45(1):55–9.
7. Gau J, Burnside WD, Gilreath M. Chebyshev multilevel absorber design concept. IEEE Trans On antennas propag. 1997;45(8):1286–93.
8. Hou XY. The measurement techniques and analysis of frequency selective surface. Chinese J Sci Instrum. 2008;29(2):36–9 (In Chinese).

Chapter 125
A Novel Planar Elliptical Dipole Antenna for Ultra-Wideband Application

Manyuan Dai, Ge Gao, Qi Wu and Hu Yang

Abstract The printed dipole antenna has been investigated to be an effective radiator for ultra-wideband (UWB) communication and radar systems. In order to broaden the impedance bandwidth of the normal planar elliptical dipole antenna, a novel planar elliptical dipole antenna, designed by means of loading balun transformer and slotting on elliptical dipole elements, is proposed. The simulation results show that the return loss of the modified antenna has been almost less than −10 dB in the frequency from 2.2 GHz to 13 GHz. Compared with the normal elliptical dipole antenna, the simulated results demonstrate that the modified elliptical dipole antenna has better directional characteristics and gain performance.

Keywords UWB (ultra-wideband) communication and radar · Planar elliptical dipole antenna · Balun transformer

125.1 Introduction

With the development of ultra-wideband (UWB) communication and radar systems, a higher performance of UWB antenna is required now. The US Department of Defense has already made the UWB and high-range resolution radar as a key technology [1]. Antenna, regarded as the component of sending and receiving electromagnetic waves, is an essential part of radar system. Synthetic aperture radar (SAR) would get higher range resolution through UWB antennas [2]. The Federal Communications Commission (FCC) in the USA certified the use of UWB devices operating in the frequency range 3.1–10.6 GHz in 2002 [3]. When the ratio of the upper frequency and lower frequency is more than one octave, it is generally considered as UWB.

Traditional UWB antennas are categorized as follows: (1) wideband dipole antenna, such as discone antenna, sleeve antenna; (2) frequency-independent antenna,

M. Dai (✉) · G. Gao · Q. Wu · H. Yang
Electronic Science and Engineering Department, National University of Defense Technology, 410073 Changsha, China
e-mail: nudt_dai@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_125

which has a self-similar structure, making it broadband characteristics naturally, such as biconical antenna, helical antenna, log periodic antenna, and planar spiral antenna; (3) transverse electromagnetic (TEM) horn antenna, which has the advantages of high gain, narrow beam, widely used in point to point communication [4]. Compared with other UWB antennas, the UWB planar dipole antenna has many advantages in easy processing, light weight, low cost, and miniaturization [5].

Elliptical dipole antenna was first investigated to have good radiation performances by N. P. Agrawall et al. in 1998 [6]. Hans Gregory Schantz achieved the bandwidth of planar elliptical dipole antenna to three octaves [7]. Hakki Nazli et al. enhanced the bandwidth of planar elliptical dipole antenna by way of loading two elliptical slots [8]. Zhen Tu et al. researched a cavity-backed elliptical dipole antenna to enhance radiation patterns [9].

The organization of this chapter is as follows. In Sect. 125.2, the parameters of the proposed antenna are introduced. In Sect. 125.3, the simulated results are presented to verify the antenna performance. Finally, the conclusions of this article are given.

125.2 Antenna Design and Geometry

The geometry of the normal planar elliptical dipole antenna is shown in Fig. 125.1a. The antenna is printed on a Rogers RO4003 substrate with relative permittivity of 3.55 and thickness of 1 mm. The minor axis of normal planar elliptical dipole antenna is about 0.3 λ_l (λ_l is the longest operation wavelength), and the major axis is about 1.25–1.75 times of the minor axis [7]. Considering the lowest operation frequency of 2.2 GHz, the minor axis of dipole elements is 4.2 cm, and the major axis of dipole elements is 5.3 cm (a = 5.3 cm, b = 4.2 cm).

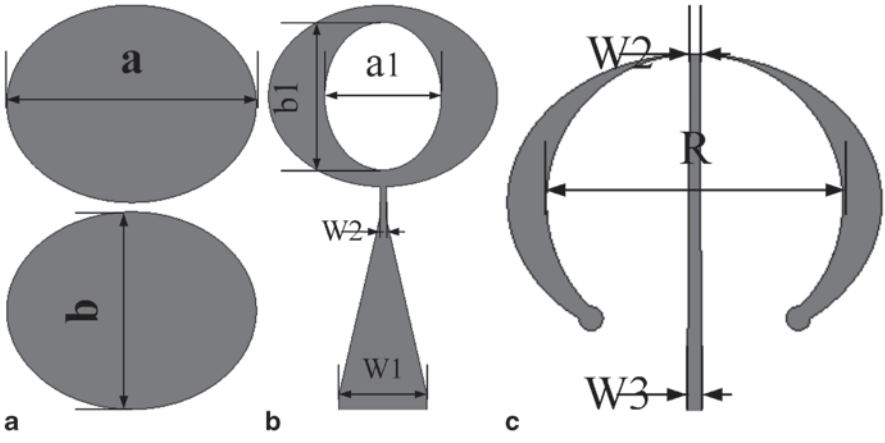


Fig. 125.1 The geometry of planar elliptical dipole antenna, (a) the normal antenna, (b) front of the modified antenna, (c) back of the modified antenna

Figure 125.1b and c shows the front and back of the modified planar elliptical dipole antenna. The outer contour of the modified antenna is the same as that of the normal antenna. In order to achieve UWB feed, a UWB gradient balun transformer is studied and designed, whose feed line is inserted into the circular slot on the back element. The circular slot is cut symmetrically with the radius $R = 7.3$ cm and the balun transformer is printed with $W1 = 20.5$ mm, $W2 = 1.43$ mm, $W3 = 2.05$ mm. As currents mainly distribute along the edge of antenna, properly slotting is also a good way to enhance the bandwidth. With loading UWB balun transformer proposed [10], the method of properly slotting on the front element is introduced in this chapter and a better radiation performance is achieved. The size of the slot on the front element is comprehensively optimized with major axis $b1 = 17.12$ mm, minor axis $a1 = 13.5$ mm, while the slot is in the center of the front element.

125.3 Antenna Performance

The antenna is simulated by using computer simulation technology (CST) microwave studio. The comparison of return loss of the normal and modified planar elliptical dipole antenna is shown in Fig. 125.2.

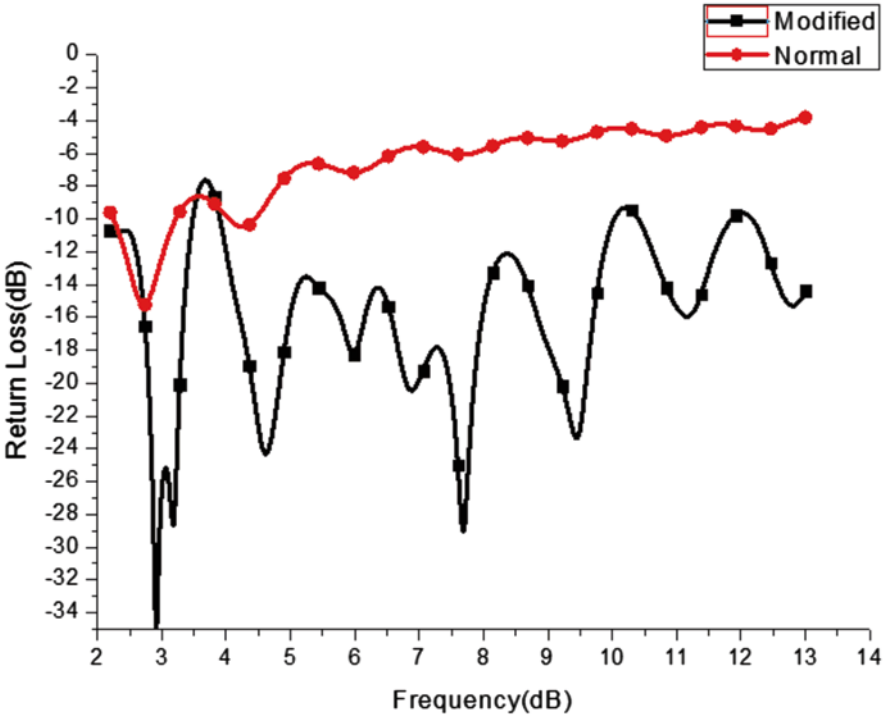


Fig. 125.2 The comparison of return loss

Compared with normal planar elliptical dipole antenna, the return loss of modified planar elliptical dipole antenna improves obviously. The return loss of the normal antenna is less than −10 dB only in the frequency from 2.2 GHz to 3.2 GHz, while the return loss of the modified antenna is less than −10 dB in most frequency from 2.2 GHz to 13 GHz, except for the frequency from 3.48 GHz to 3.9 GHz where the return loss is less than −7.3 dB. At the same time, it could be verified that by the method of slotting properly, the bandwidth of the modified antenna is further enhanced and significantly wider than the bandwidth, which is only from 2.5 GHz to 6 GHz [10].

The Figs. 125.3, 125.4, and 125.5 present E-plane and H-plane radiation patterns for certain frequencies. In the low frequency range, these antennas behave like dipoles. As sidelobes appearing with the frequency increase, the omnidirectional characteristics of the antenna deteriorate gradually. However, the modified planar elliptical dipole antenna has smaller deterioration than the normal planar elliptical dipole antenna in omnidirection (Fig. 125.6).

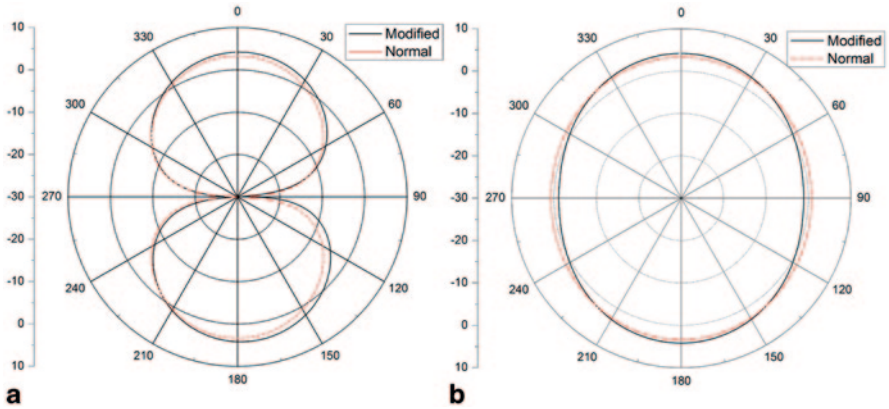


Fig. 125.3 Radiation patterns when f = 2.5 GHz (a) E-plane (b) H-plane

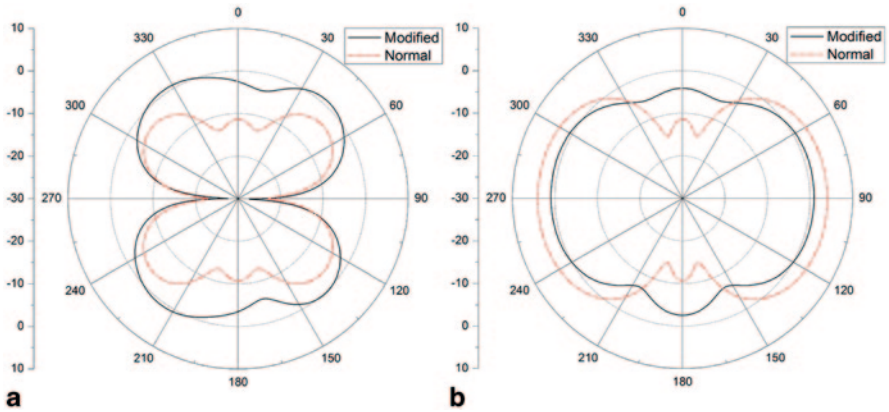


Fig. 125.4 Radiation patterns when f = 6 GHz (a) E-plane (b) H-plane

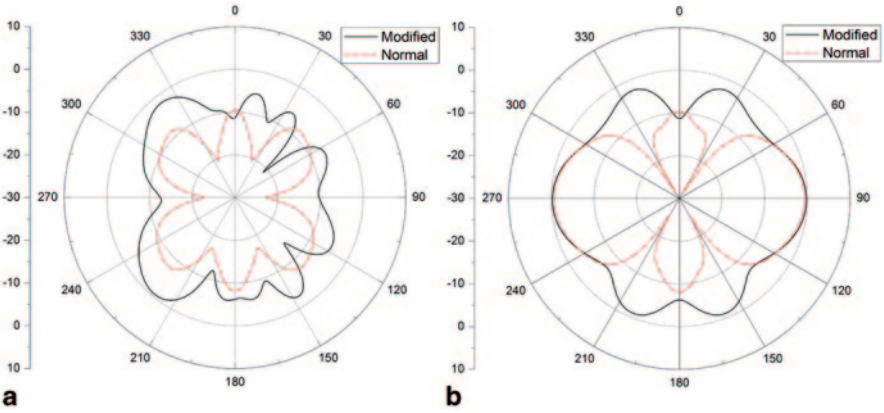


Fig. 125.5 Radiation patterns when f = 11 GHz (a) E-plane (b) H-plane

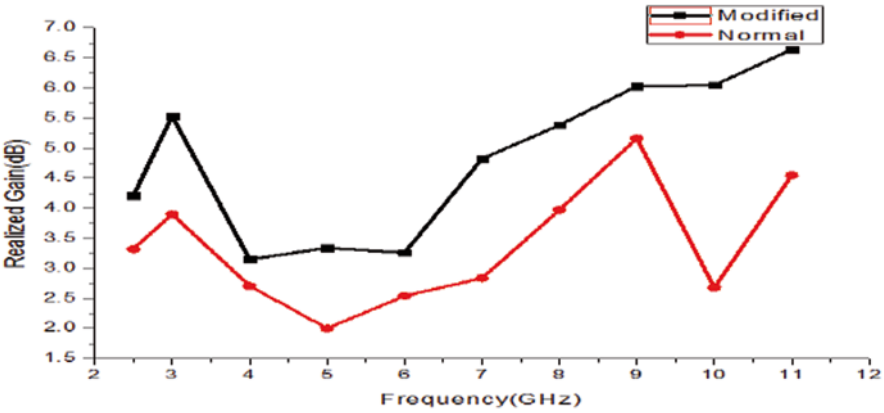


Fig. 125.6 The comparison of realized gain of maximum radiation direction

As it can be seen, the realized gain of the modified planar elliptical dipole antenna is higher than the normal antenna. With the frequency increasing, the maximum realized gain of the modified antenna is also increasing, because the omnidirectional characteristics of the antenna deteriorate gradually, which results in strong directional characteristics and narrow beams.

125.4 Conclusion

Nowadays, the rapid development of UWB communication and radar system put forward higher requirements for UWB antenna. Compared with the normal planar elliptical dipole antenna, a novel planar elliptical dipole antenna with balun trans-

former is proposed, which has been verified to broaden operation bandwidth to six octaves and achieve a higher realized gain. And the modified antenna maintains advantages in arraying and low cost, which could be widely applied to UWB communication and radar systems.

References

1. Haiying Li, Ruliang Yang. Development, state-of-the-art and applications of ultra-wideband radar. Remote Sens Technol Appl. 2001;16(3):178–82 (In Chinese).
2. Chengbo Z. Synthetic aperture radar theory, system analysis and application. Beijing: Science Press; 1989. pp. 176–9 (In Chinese).
3. Federal Communications Commission (FCC). Revision of part 14 of the commission's rules-regarding ultra wideband transmission systems. FCC 02–48; 2002.
4. Schantz HG. A brief history of UWB antennas. Aerosp Electron Syst Mag. 2004;19(4):22–6.
5. Li C, Yang S, Zhang C. Review of ultra-wide band pulse antenna Chinese. J Radio Sci. 2008;23(5):1003–8 (In Chinese).
6. Schantz HG. The art and science of ultrawideband antennas. Boston: Artech House; 2005. pp. 1–331.
7. Schantz HG. Planar elliptical element ultra-wideband dipole antennas. IEEE Antenna Propagat Soc Int Symp. 2002;3(6):44–7.
8. Nazli H, Bicak E, Turetken B, Sezgin M. An improved design of planar elliptical dipole antenna for UWB applications. IEEE Antennas Wirel Propagat Lett. 2010;9(3):264–7.
9. Zhen Tu, Dongfang Zhou, Guangqiu Zhang, Feng Xing, Xue Lei, Dewei Zhang. A wideband cavity-backed elliptical printed dipole antenna with enhanced radiation patterns. IEEE Antennas Wirel Propagat Lett. 2013;12(11):1610–3.
10. Schantz HG. Bottom fed planar elliptical UWB antennas. UWB systems and technologies; 2003 IEEE Conference on, 16–19 Nov. 2003, IEEE, pp. 219–23.

Chapter 126
An Anticollision Protocol of Adaptive Tree Slot Without Tag Estimation

Qin Zhu and Haifeng Wu

Abstract In order to improve the throughput of tags identification in the system of radio frequency identification (RFID), the anticollision protocol of adaptive tree slot is proposed, and that is a tag anticollision algorithm in view of the Q algorithm. In this protocol, the rest of the tags will be waiting until those that are in the first collision slot are decomposed with the method of the binary tree. Besides, the tags in the rest of the collision slots will be resolved by the tree slot. The advantage of the proposed protocol lies in the fact that it does not need the prior information about tags, which can reduce the complexity of the system; furthermore, the system with high identification efficiency can be kept, and the change of the tag quantity cannot influence it. It indicates in our simulations that the identification efficiency of the system does not fluctuate much when the number of tags is between 40 and 500; moreover, when the tag quantity does not approach to the original frame length, the identification efficiency of the system is higher than the tree slotted ALOHA (TSA) algorithm and fast k-slot selection (FKSS) protocol.

Keywords RFID · Anti-collision · Identification efficiency · Tree slotted Aloha · Adaptive tree slot

126.1 Introduction

Radio frequency identification (RFID) plays an important role in the technology of Internet of things, the most important problem of RFID system is the tag anticollision [1–5]. At present, the tag anticollision has been solved based on two aspects, which are respectively based on the ALOHA protocol [6] and the type of tree anticollision algorithm [7]. In every slot, the slot is defined as a success slot if only one

Q. Zhu (✉) · H. Wu
School of Electrical and Information Technology, Yunnan University of Nationalities, 650500 Kunming, China
e-mail: whf5469@gmail.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control*, Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_126

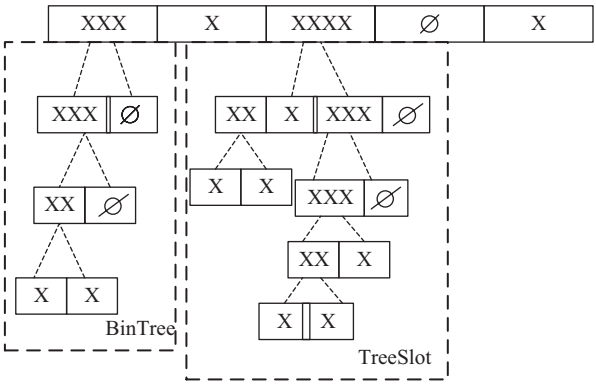


Fig. 126.1 Diagram of the process of applied type system (ATS) algorithm

tag exists, and it is a collision slot if there are more than two tags. It is an idle slot if there is no tag. To enhance the efficiency of the system, the thought of the tree protocol is introduced into the ALOHA protocol by someone, which develops mixed anticollision protocols, and the tree slotted ALOHA (TSA) [8] is the one with high throughput.

Fast k-slot selection (FKSS) protocol refers to the fact that the length of the frame is regulated to attain the optimum by means of judging the first k-slots. Its advantage is that it does not need the prior information about tags, but the highest throughput is gotten only when the length of the frame approaches to the tag quantity. Q algorithm [9–10] adopts the self-adaptive thoughts, and it regulates the length of the frame based on the present slot's condition. Like FKSS algorithm, Q algorithm also does not need the prior information about tags; however, the throughput of Q algorithm can only attain 0.34. Compared with the above two algorithms, the disadvantage of the TSA protocol is that the change of the tag quantity influences the system throughput. Generally, the tag quantity in the system is unknown; however, the high identification efficiency of the system is achieved when the length of the frame corresponds to the tag quantity, which needs the prior information about tags; therefore, it increases its complexity and decreases its application.

An anticollision protocol of adaptive tree slot without tag estimation on the basis of Q algorithm is presented in this chapter. Its thought is shown in Fig. 126.1. According to the present status of the slot, the frame length is adjusted dynamically. The tags in the first collision slot are decomposed by a binary tree in order to avoid wasting the slots; the length of the frame will be reset under the condition of the remaining collision slots, and those collision tags will be decomposed by the tree slot.

The main body of the chapter is arranged as explained below. The extant problem is introduced in Part II, and the fundamental thought of the TSA protocol is also presented in this part; the adaptive tree slot and its elementary process are proposed in Part III. The performance of the protocol based on the analysis of experimental simulation results is put forward in Part IV; the conclusion is made in Part V.

126.2 Extant Problems

The thought of the TSA protocol is that the frame is divided into some slots, and tags are randomly distributed to those slots. If there is a collision in the h^{th} slot, which occurs in the frame of the $m-1^{\text{th}}$ layer, the collision tags will need the frame length that is Y_m^h in the h^{th} slot of the next layer to decompose, and it is supposed to be equivalent between the value of Y_m^h and the population of the collision tags X_{m-1}^h in the h^{th} slot of the upper layer, which is:

$$Y_m^h = X_{m-1}^h, m > 0, h \geq 0. \tag{126.1}$$

Nevertheless, the number of the collision tags is unknown to the reader; therefore, the TSA protocol determines according to the following way: if a_{m-1} tags are distributed in the frame of the $m-1^{\text{th}}$ layer, there are $b_{1,m-1}$ success slots and $b_{t,m-1}$ collision slots, then the collision tags of the upper layer will be identified in the frame of the m^{th} layer by using the length of Y_m. Y_m is determined by Eq. (126.2), and it can be applied in any slots.

$$Y_m = \left\lfloor \frac{a_{m-1} - b_{1,m-1}}{b_{t,m-1}} \right\rfloor, m > 0, h \geq 0. \tag{126.2}$$

The identification efficiency comparison among TSA, Q algorithm, and FKSS algorithm is shown in Fig. 126.2, where the initial value of the frame length of Q

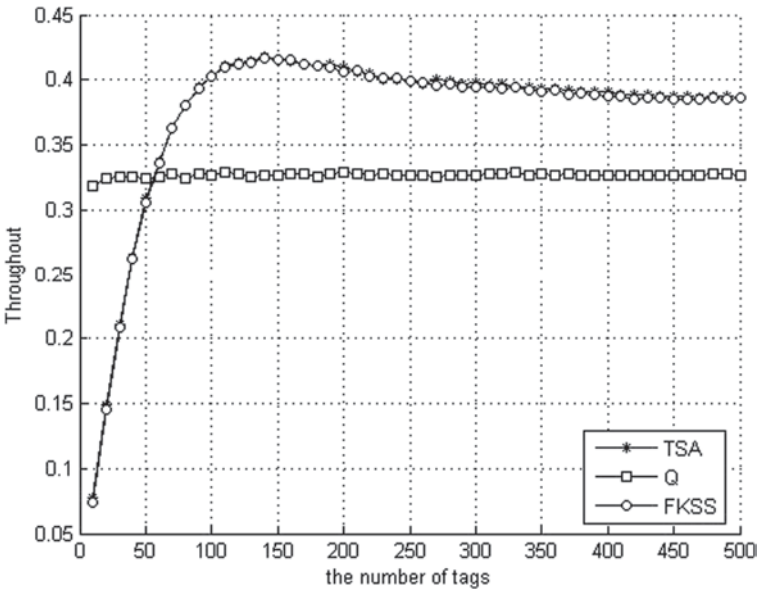


Fig. 126.2 Identification efficiency comparison among tree slotted ALOHA (*TSA*), Q algorithm, and fast k-slot selection (*FKSS*) algorithm

algorithm is $Y_0=16$. Those of TSA and FKSS algorithm are $Y_0=128$, and the number of tags is 500. As it is shown, only if the tag quantity approaches the original frame length, the maximum of the throughput of TSA and FKSS algorithm is achieved; if the difference between the tag quantity and the original frame length is big, then the throughput of TSA and FKSS algorithm decline distinctly. Besides, the maximum of the throughput of Q algorithm is only 0.34, which is below the TSA and FKSS algorithm.

126.3 Anticollision Protocol of Adaptive Tree Slot

The adaptive tree slot protocol adjusts the frame length in a self-adaptive way by judging the state of the current slot, which makes it meet certain rules between the tag quantity and the length of the frame, and collision tags will be decomposed by using the tree slot in the current slot to achieve a higher efficiency.

The adaptive tree slot protocol is based on the Q algorithm, and its specific operation is shown below: If too many collision slots occur in a frame, the frame will be finished by the reader in advance, and a longer frame will be transmitted. On the contrary, a shorter frame will be transmitted because of too many idle slots. The above process will be implemented through three parameters: Q, Qfp, and C. The value of $Y=2^Q$ is the frame length, and the value of Q changes dynamically, and its value is $round(Q_{fp})$. Whether a new frame is opened depends on whether Q is equal to $round(Q_{fp})$ judged by the reader. In the adaptive tree slot protocol, the parameters adopt the rule in the Q algorithm, and the value of 4.0 is set to the original value of Q; in addition, the range of the step value is $0.1 \le C \le 0.5$.

The difference between the Q algorithm and the proposed protocol is that in the Q algorithm, if $round(Q_{fp})=Q$, the tag is identified and goes onto the next slot; however, in the proposed algorithm, if $round(Q_{fp})=Q$ and the collision occurs, with the method of the binary tree, the tags that lie in the first collision slot is decomposed, and the rest of them will wait to be identified until the tags in the collision slot finish decomposing. As for other collision slots below, the frame length will be reset, and the tags will be decomposed by using tree slot.

126.4 Simulation Result

The simulation result will be presented in this portion, and the average result after 500 independent experiments by using the method of Monte Carlo is gotten.

Figure 126.3 shows the throughput of the tag identification of different algorithms. The tag quantity is between 10 and 500, and the initial value of the frame length is $Y_0=128$. As we can see, the throughputs of TSA algorithm and FKSS algorithm are high; however, the highest throughputs are both achieved when the length of the frame approaches the tag quantity. The Q algorithm is with the lowest throughput, whose maximum is about 0.34. The ATS algorithm is of high throughput; by comparison, it is with steady change. When the tag quantity is around 40–500, the efficiency of the tags identification is approximate to a straight line; besides, when

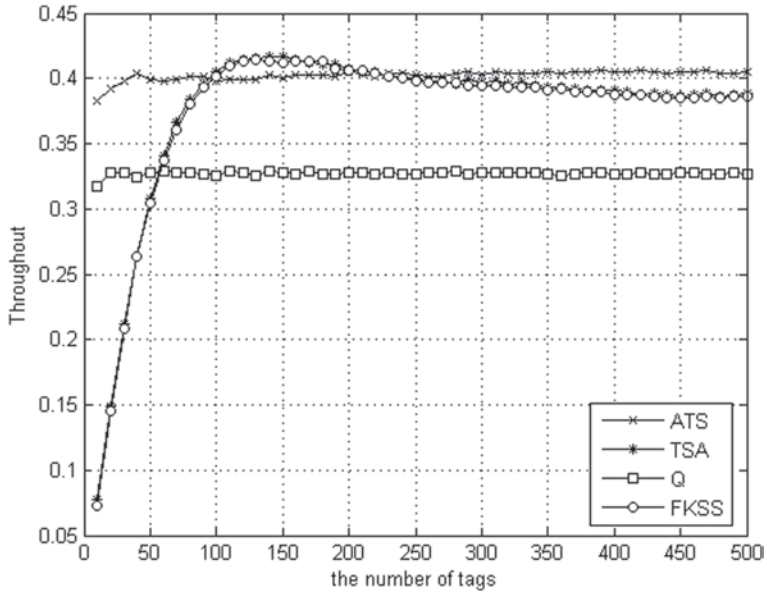


Fig. 126.3 Throughput of the tags identification of different algorithms. *ATS* applied type system, *TSA* tree slotted ALOHA, *FKSS* fast k-slot selection

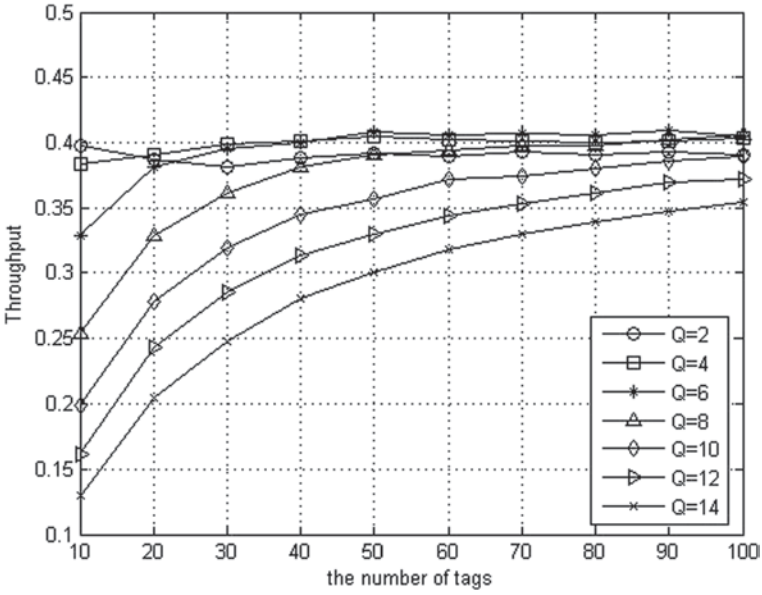


Fig. 126.4 Throughput of ATS algorithm with diverse Qs

the tag quantity does not approach the original frame, the curve of the throughput of the ATS protocol is above TSA algorithm.

Figure 126.4 shows the throughput of the ATS algorithm with diverse Qs, and the tag quantity is between 10 and 500. As it is shown, if the value of 4 is set to the

original Q value, it is relatively reasonable, and it can get a satisfying result independent of the number of tags.

126.5 Conclusion

In this chapter, the anticollision of adaptive tree slot with adaptive frame length is proposed. The proposed algorithm can make a relative relation between the length of the frame and the tag quantity; besides, the tag quantity cannot influence the throughput of the system. As far as the results of the simulation are concerned, the anticollision of adaptive tree slot proposed is of high throughput of the system, and it does not need prior information about tags, which can reduce the complexity of the system.

Acknowledgments This work was supported partially by the National Natural Science Foundation of China under No. 61262091 Grant, the Applied and Basic Research Foundation of Yunnan Province under No. 2011FB083 Grant, the Major Special Project of Scientific Research Foundation of Yunnan Provincial Department of Education under No. ZD2011009 Grant, a grant from Innovative Research Team in Yunnan University of Nationalities, and the 17th batch of Young and Middle-Aged Leaders in Academic and Technical Reserved Talents Project of Yunnan Province.

References

1. Shin D, Sun P, Yen D, Huang S. Taxonomy and survey of RFID anti-collision protocols. Comput Commun. 2006;29(11):2150–66.
2. Wu H, Zeng Y. Efficient framed slotted aloha protocol RFID tag anti-collision. IEEE Trans Autom Sci Eng. 2011;8(3):581–8.
3. Klair D, Chin K, Raad R. A survey and tutorial of RFID anti-collision protocols, communication survey & turials, IEEE, third quarter; 2010. p. 400–21.
4. Wang L, Liu H. A novel anti-collision algorithm for EPC gen2 RFID systems. In Proceeds of IEEE International Conference on system, man and cybernetic. Hammamet, Tunisia: IEEE; 2006. p. 7–6.
5. Maguire Y, Pappu R. An optimal Q-algorithm for the ISO 18000-6 C RFID protocol. IEEE Trans Autom Sci Eng. 2009;6(1):16–24.
6. Lee S, Joo S, Lee C. An enhanced dynamic framed ALOHA algorithm for RFID tag identification. Proceedings International Conference mobile and ubiquitous systems: networking and services, USA: IEEE; 2005. p. 1–5.
7. EPC. Radio-frequency identity protocols class-I generation-2 UHF RFID protocol for communications at 860 Mhz–960 Mhz version 1.0.9 2004.
8. Bonuccelli M, Lonetti F, Martelli F. Tree slotted ALOHA: a new protocol for tag identification in RFID networks//Proc of Int Symp on a world of wireless, mobile and multimedia networks. Piscataway, NJ: IEEE; 2006. p. 1–6.
9. Lee D, Kim K, Lee W. Q-Algorithm: an enhanced RFID tag collision arbitration algorithm. Comput Sci. 2007;4611:23–32.
10. EPCTM. Radio-frequency identify protocols class-1 generation-2 UHF RFID protocol for communications at 860 MHz–960 MHz Version 1.1.0 Draft1, EPCglobal Inc, 2005.

Chapter 127
Infrared Small Target Detection Using Two-Dimensional Least Mean Square Filter Based on Neighborhood Information

Lili Wan and Min Wang

Abstract This chapter proposes an infrared small target detection algorithm using two-dimensional least mean square (TDLMS) filter based on neighborhood information. The structure of the TDLMS filter and prediction method is improved to make full use of neighborhood information of the predicted pixel, and the background is predicted using a nonlinear step adjustment method to improve the prediction accuracy. Experimental results show that the background can be effectively suppressed, and the detection rate of infrared small target is improved if the background is predicted by the TDLMS filter based on neighborhood information.

Keywords TDLMS filter · Neighborhood information · Background prediction · Infrared small target

127.1 Introduction

Infrared target detection technology has broad application prospects, and it is widely applied in military and civilian fields. In the military field, the infrared target detection technology is a key technique in precision guidance and warning; besides, it can be used in industry, agriculture, medicine, and transportation in the civilian field. In terms of industry, the infrared small target detection technology can be used for the purpose of industrial control, robot vision, autonomous navigation, etc. In terms of medicine, it can be applied in biological tissue motion analysis, and it can also be used for traffic management, transport flow control in terms of transportation; however, due to the strong noise and interferences, the infrared target detection becomes difficult especially when the target is small. The target is easy to be submerged with the background because of the low signal-to-noise ratio (SNR);

M. Wang (✉) · L. Wan
School of Automation, Huazhong University of Science and Technology,
430074 Wuhan, China
e-mail: wm526@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_127

therefore, the detection of the infrared small target is a challenging task, and the research is of great theoretical significance and practical value.

The background prediction is an important research of infrared small target detection with many achievements having been obtained, for example, the background prediction method based on the fixed weight linear low-pass filter is simple, fast, and easy to implement, but it has some limitations due to its fixed weights [1–2]. The median filtering is an important method for smoothing background, which can attenuate random noise and retain edge information of the object at the same time [3–4]. The top-hat transform is also a practical infrared image background prediction method, which features good real-time performance, but it is mainly applied for the procession under a smooth and slowly varying background [5–7]. Two-dimensional least mean square (TDLMS) filter is an effective self-adaptive filter, which shows promising results in the infrared small target detection [8]; however, it still needs to be further improved. Many researchers have been devoted to improving the performance of the TDLMS filter and have already made much process [9–11].

Tae-Wuk Bae et al. proposed an infrared small target detection method by using the edge directional 2D LMS filter [11]. The edge direction information of surrounding prediction pixels was used to predict the background pixels excluding small targets. This method has a better performance in chaotic regions than TDLMS filter based on neighborhood analysis [10], but it still cannot flexibly respond to changes in the background because only four directions are evaluated.

127.2 TDLMS Filter

The TDLMS filter was proposed by Hadhoud and Thomas in 1988, as the expansion of one-dimensional LMS adaptive filter [8]. Figure 127.1 shows the structure of the TDLMS filter. The size of the reference input image (X) is $M \times N$, and the size of the filter window is $P \times P$. The output of the filter is given by Formula (127.1).

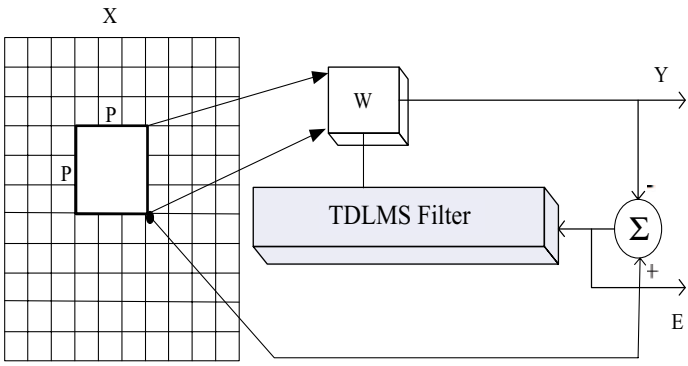


Fig. 127.1 Basic structure of two-dimensional least mean square (*TDLMS*) filter

$$Y(m,n)=\sum_{l=0}^{p-1}\sum_{k=0}^{p-1}W_j(l,k)X(m-l,n-k), m=(0,1,\cdots,M-1), n=(0,1,\cdots,N-1), \tag{127.1}$$

where W_j is the weight matrix at the j th iteration. If the image is scanned in a left to right manner, then $j=m\times N+n$. The weight matrix, W_j, is updated with the steepest decent method, which is given by

$$W_{j+1}(l,k)=W_j(l,k)+\mu e_j X(m-l,n-k), \tag{127.2}$$

where e_j is the prediction error, which is calculated by $e_j=D(m,n)-Y(m,n)$. μ is the step size, and it is closely related to the convergence rate [12]. When the filter is stable, a larger μ, the filter can respond timely to changes, thus both the background and targets will be predicted precisely. The subtraction image, which is obtained by subtracting the predicted background from the original image, will have a lower SNR; thus, it is difficult to detect the small targets. On the other hand, a smaller μ makes the weights converge slower, and both the target regions and the background regions will be ignored in great degree, so the background prediction accuracy will decrease and some pseudo targets might appear.

The TDLMS filter is used as a background predictor in the infrared small target detection field. When a part of the background is covered by the filter window, the algorithm tends to converge after several iterations, and the background can be predicted because the infrared image background has strong correlation. When the filter window comes to a small target, the algorithm cannot achieve convergence before the filter window moves to the background region because the correlation of the target and the background is small; therefore, the output of the filter is the continuation of the background.

127.3 Proposed TDLMS Filter Based on Neighborhood Information

127.3.1 Structure of Proposed TDLMS Filter

In this chapter, an improvement to edge directional 2D LMS filter has been made. The structure of proposed TDLMS filter is illustrated in Fig. 127.2a. Comparing with the edge directional 2D LMS filter, the output block has been improved, i.e., the output block based on edge direction is replaced with output block based on neighborhood information. The proposed TDLMS filter uses eight input blocks, namely $X_i(m,n)(i=0\cdots3,5\cdots8)$ around the predicted pixels $X(m,n)$, in which all the blocks have $P\times P$ size as shown in Fig. 127.2b. Each input block will generate the corresponding predictive value $Y_i(m,n)$, i.e.,

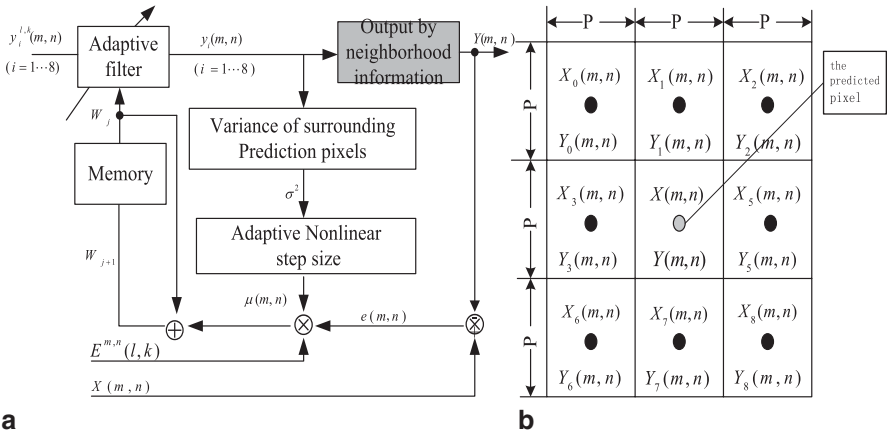


Fig. 127.2 **a** Structure of TDLMS filter based on neighborhood information. **b** Input blocks

$$Y_i(m,n)=\sum_{l=0}^{p-1}\sum_{k=0}^{p-1}W_j(l,k)X_i(l,k),\ i=0\cdots3,5\cdots8. \tag{127.3}$$

These eight input blocks share the same weight matrix, which is updated by

$$W_{j+1}(l,k)=W_j(l,k)+\mu e_j(m,n)A^{l,k}(m,n),e_j(m,n)=X_j(m,n)-Y_j(m,n). \tag{127.4}$$

127.3.2 Output Based on Neighborhood Information

In order to improve the algorithm target extraction and background suppression ability, the output of the TDLMS filter is decided according to the relationship between eight predictive values and the predicted value. When the predicted value is greater than eight predictive values, it indicates that the center of input blocks has been moved to a small target as shown in Fig. 127.3a. Under this situation, the difference between the predicted value and the gray value of small target should be large enough to ignore the small targets; therefore, the average of two smallest predictive values is the output of the TDLMS filter, which can be expressed by

$$Y(m,n)=\frac{Y_j+Y_k}{2},Y_j=\min\{Y_i(m,n)\},Y_k=\min\{Y_i(m,n),i\neq j\},i=1\cdots3,5\cdots8. \tag{127.5}$$

If the predicted value is not greater than eight predictive values, it means the center of input blocks may cover the background regions as shown in Fig. 127.3b. In order to accurately predict the background, the predictive value needs to be as close as possible to the predicted value. Then, the proposed method calculates the absolute

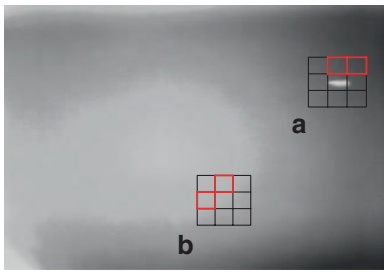


Fig. 127.3 Input blocks in different areas

difference between the predictive values and the currently processed pixel value $X(m,n)$, and the formula can be represented as below:

$$d_i(m,n) = \left|Y_i(m,n) - X(m,n)\right|, (i = 1\cdots 3, 5\cdots 8). \tag{127.6}$$

The filter output $Y(m,n)$ based on neighborhood information is given by

$$Y(m,n) = \frac{d_j + d_k}{2} d_j = \min\{d_i(m,n)\}, d_k = \min\{d_i(m,n), i \neq j\}, i = 1\cdots 3, 5\cdots 8. \tag{127.7}$$

If the filter output is calculated by Formula (127.5) or (127.7), we know the corresponding subscripts j,k. Then, the block $A^{l,k}(m,n)$ in Eq. (127.4) is updated by

$$A^{l,k}(m,n) = \frac{X_j^{l,k}(m,n) + X_k^{l,k}(m,n)}{2}. \tag{127.8}$$

Figure 127.3 shows two situations when the filter window covers the target region and the background region. In the small target region of Fig. 127.3a, the filter output of TDLMS filter is the average of the two smallest predictive values ($Y_1(m,n), Y_2(m,n)$ in Fig. 127.2b) corresponding to two red blocks in Fig. 127.3a. If the output of TDLMS filter is based on edge direction, the filter output will be the average of $Y_3(m,n), Y_5(m,n)$. The difference between the filter output and the small target gray value will become smaller so that the small target might become too dim to detect or will even disappear. In the edge region in Fig. 127.3b, the filter output of TDLMS filter uses two predictive values $Y_1(m,n), Y_3(m,n)$ in Fig. 127.2b can predict the background more accurately than the filter output using $Y_2(m,n), Y_6(m,n)$, which will be selected by edge directional 2D LMS filter, because the average of $Y_1(m,n), Y_3(m,n)$ is closer to the predicted value than the average of $Y_2(m,n), Y_6(m,n)$. There would be fewer false alarms, and the prediction accuracy will be improved.

127.3.3 Adaptive Nonlinear Step Size

In this chapter, adaptive region-dependent nonlinear step size will be adopted [11]. The step size will be updated according to the variance of predictive values excluding the maximum and minimum among $Y_i(m,n)$, and it is updated by

$$\mu(m,n) = \mu_{\min} + \frac{\mu_{\max}}{1+e^{-\log\sigma^2_{g(m,n)}}}, \tag{127.9}$$

where $\mu_{\max}, \mu_{\min}$ are the maximum and minimum μ. $\sigma^2_{g(m,n)}$ is the variance of the set

$$g(m,n) = \{Y_i(m,n) \mid i = 1\cdots 3, 5\cdots 8, \min(Y_i(m,n)), \max(Y_i(m,n)) \notin g(m,n)\}.$$

127.4 Experimental Results

Computer simulations are conducted on three video sequences, which contain one or more small dim flying targets. In these experiments, we choose $\mu_{\max} = 1\times 10^{-10}$, $\mu_{\min} = 1\times 10^{-11}$, and $P\times P = 3\times 3$. Figure 127.4 shows the detection results by using the proposed algorithm. The detection results will be obtained by binarizing the subtracted image. The binarization image will be obtained through adaptive threshold segmentation based on minimum error [13].

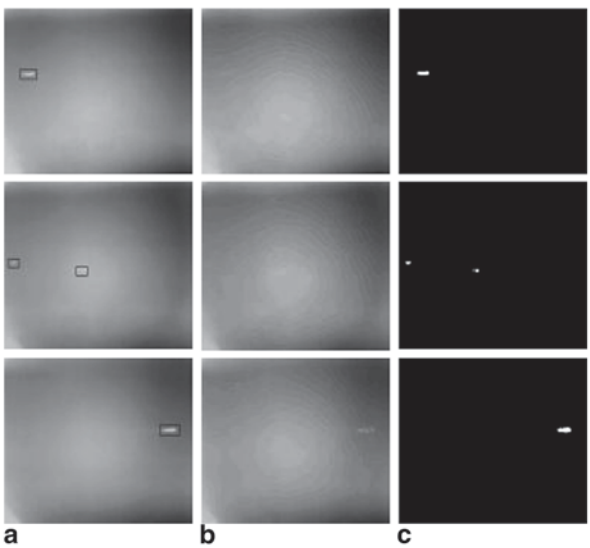


Fig. 127.4 Detection results of infrared small targets. **a** Original images. **b** Background prediction images. **c** Result images

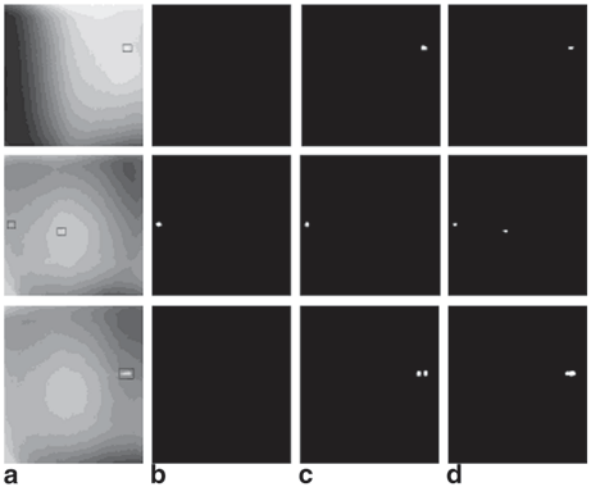


Fig. 127.5 Detection results of three methods. **a** Original images. **b** Detection results using the basic filter. **c** Detection results using Tae-Wuk Bae's filter. **d** Detection results using the proposed filter

Figure 127.4 shows that the backgrounds of the original images have been estimated precisely, and the targets are ignored or even disappeared in the predicted images. Even if the contrast ratio between the target and the background is very low, the small targets could be detected by the proposed algorithm, which also has a good performance in multi-target detection. In the experiment, we can see that the proposed algorithm has good performances.

In order to fully illustrate the performance of the proposed algorithm, we make a comparison among the detection results using the proposed TDLMS filter, edge directional 2D LMS filter, and the basic TDLMS filter. Figure 127.5 shows the detection results of three different methods. From the detection results of the first test image, we can see that the proposed TDLMS filter and Tae-Wuk Bae's edge directional 2D LMS filter have a better performance under low contrast conditions. The detection results of the second test image illustrates the proposed TDLMS filter has a better performance than the other two filters in multi-target detection. The detection results of the third test image indicates the shortage of the edge directional 2D LMS filter, which might result in the incomplete extraction of the target, while the proposed filter can extract the targets integrally.

To compare these three methods quantitatively, two metrics, the detection rate (DR) and the false alarm rate (FAR), are employed and defined as below:

$$DR = {T_n}/{T_N}, \quad FAR = {F_e}/{F_N} \tag{127.10}$$

Where T_n is the number of targets which is detected correctly, T_N is the number of targets in video, F_e is the number of frames which has false alarm, and F_N is the total number of frames of the video. The statistical results are listed in Table 127.1.

The statistical results reveal that the proposed algorithm has effective performance in both DR and FAR. The experimental results prove that the proposed algo-

Table 127.1 Performance comparison of three filters

	Video 1		*Video 2*		*Video 3*	
	DR (%)	FAR (%)	DR (%)	FAR (%)	DR (%)	FAR (%)
Basic TDLMS filter	86.7	0.579	78.6	0.143	90.65	0.521
Edge directional 2D LMS filter	93.45	6.94	87.3	4.29	95.32	3.75
Proposed TDLMS filter	97.17	0.46	90.5	0.286	96.01	2.82

DR detection rate, *FAR* false alarm rate, *TDLMS* two-dimensional least mean square

rithm can improve detection ability and attain best balance between DR and FAR. From the comparison, we can clearly see that the proposed algorithm has significant advantages in DR. Subjective and objective comparison results reveal that the proposed algorithm has better target detection and background prediction ability.

127.5 Conclusion

The proposed algorithm can respond flexibly to the changes in the edge region and predict the background exactly, so it can make small targets in subtraction image easy to be distinguished. Experimental results demonstrate that the proposed method has good target detection and background prediction ability in the infrared small target detection field. Compared with the other two algorithms, the performance of the proposed algorithm is better, especially in DR.

Acknowledgments The study was supported by the National Natural Science Foundation of China (61370180).

References

1. Xu J. Research on the detection of small and dim targets in infrared images. Xi'an: Xi'an University of Electronic Science and Technology PhD thesis. (In Chinese); 2003.
2. Scharf LL, Friedlander B. Matched subspace detectors. IEEE Trans Signal Proc. 1994;42(8):2146–56.
3. Yang W, Shen Z. Small target detection and preprocessing technology in infrared image sequences. Infrared Laser Eng. 1998;27(1):23–28. (In Chinese)
4. Zhang W, Wu J, Zhu Z, Li H, Wang Q. Application of improved crossing based on median filter in image processing of IR image. Electron Opt Contro 2006;1(1):83–6. (In Chinese)
5. Wang Y, Zheng Q, Zhan J. Real-time detection of small target in IR grey image based on mathematical morphology. Infrared Laser Eng. 2003;30(1):28–31. (In Chinese)
6. Sun W, Xia L. Infrared target segmentation algorithm based on morphological method. J Infrared Millim Waves. 2004;23(3):233–6. (In Chinese)
7. Ye B, Peng J. Small target detection method based on morphology top-hat operator. J Image Graph. 2002;7(7):15–9. (In Chinese)

8. Hadhoud MM**,** Thomas DW. The two-dimensional adaptive LMS(TDLMS) algorithm. IEEE Trans Circuits Syst. 1988;35(5):485–94.
9. Ffrench PA, Zeidler JR, Ku WH. Enhanced detectability of small objects in correlated clutter using an improved 2-D adaptive lattice algorithm. IEEE Trans Image Process. 1997;6(3):383–97.
10. Cao Y, Liu R, Yang J. Small target detection using two-dimensional least mean square (TDLMS) filter based on neighbor analysis. Infrared Millim Waves. 2008;29(2):188–200.
11. Bae TW, Zhang F, Kweon IS. Edge directional 2D LMS filter for infrared small target detection. Infrared Phys Technol. 2012;55(1):137–45.
12. Zhao S, Man Z, Jones DL, Khoo S. A Variable Step-Size Transform-Domain LMS Algorithm Based on Minimum Mean-Square Deviation for Autoregressive Process. Proceedings of the 8th IEEE Conference on Industrial Electronics and Applications. p. 968–71; 2013.
13. Cho S, Haralick R, Yi S. Improvement of kittler and illingworth's minimum error thresholding. Pattern Recognit. 1989;22(5):609–17.

Chapter 128
Infrared Dim and Small Target Detection Based on the Human Visual Attention Mechanism

Yaping Deng and Min Wang

Abstract This article proposes an algorithm based on the human visual attention mechanism to solve the infrared target detection problem provided that the targets are submerged in the background. Firstly, the regions of interest (ROIs) of the image are obtained by the top-hat transform of mathematical morphology and the method of adaptive thresholding segmentation. Secondly, the image's signal-to-noise ratio (SNR) is enhanced by processing the ROIs using the difference of Gaussians (Dog) filter which has the characteristic of human vision in the scale space. Then, the points which have the local maximum SNR in the detected image can be regarded as the candidate targets. At last, considering the targets are easily submerged in the background and to prevent the targets not being detected, the algorithm proposes searching the missing targets again using the Retinex theory. Experimental results with real forward-looking infrared (FLIR) images show higher detection rate and lower false alarm rate than other methods, especially for the targets submerged in the background.

Keywords Visual attention mechanism · Top-hat · Dog · Scale space · Retinex

128.1 Introduction

The research and application of target detection are an important branch of computer vision. With the development of infrared image technology, the detection of infrared targets draws great attention to the researchers at home and abroad, especially the detection of the infrared target in the complex background.

Recently, many algorithms have been provided to deal with the detection of the infrared target. They involve single-frame detection [1–5] and multi-frame detection [6, 7]. With the further study of researchers, the algorithm of infrared target detection based on biological vision draws great attention and has good effects [8–12]. Sungho Kim et al. [11] used the filter of Laplacian of Gaussian (Log) in the scale

M. Wang (✉) · Y. Deng
School of Automation, Huazhong University of Science and Technology, 430074 Wuhan, China
e-mail: wm526@163.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_128

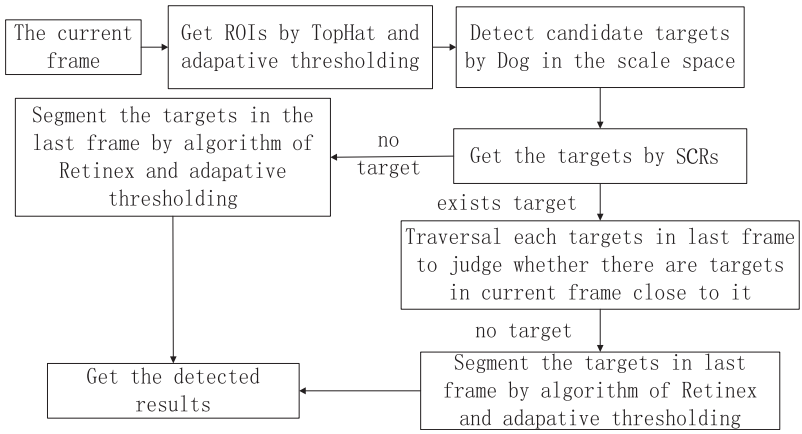


Fig. 128.1 The flowchart of the algorithm. *ROI* region of interests, *Dog* difference of Gaussians, *SCR* signal-to-clutter ratio

space to improve the signal-to-noise ratio (SNR) of the images, then Tune-Max of the signal to background contrast and the method of Tune-Max of the signal-to-clutter ratio (SCR) were used to detect the targets. Shao Xiaopeng et al. [12] firstly proposed the algorithm using the filter of Log and morphological methods to pre-process the image, then true targets could be obtained by applying local thresholding segmentation.

The algorithms above can have good effects in most of the conditions but cannot effectively detect the targets when they are submerged in the background. Considering this, this article proposes a new algorithm based on the human visual attention mechanism. Instead of using all the data in the image to detect the targets, it just focuses on small parts of the image called the regions of interest (ROIs). After processing the ROIs using the Dog filter in the scale space, the points whose SCR is a local maximum and bigger than a threshold are regarded as the targets. In order to prevent some targets are missing, the local domain of each targets which are judged to be missing is enhanced by the algorithm of multi-scale Retinex (MSR). By segmenting the enhanced local domain the missing targets can be detected. Figure 128.1 shows the flowchart of the proposed method.

128.2 The Analysis of the Algorithm

128.2.1 The Achievement of the ROI

Considering the initial process of searching the targets in the human visual system is to locate the targets in a small area of the image, this algorithm just processes the ROIs of the image. Through this, it can reduce the complexity of the process.

128.2.1.1 Top-Hat

Morphological methods have a wide application in the preprocess of the infrared target detection. This algorithm adopts the usual filter top-hat to suppress the background and enhance the contrast of the targets. The top-hat filter is defined as (128.1):

$$Tophat = A - (A^{\circ} B), \tag{128.1}$$

where $[A^{\circ} B = (A \Theta B) \oplus B$. It is called the opening operation. The opening A by B is the erosion of A by B, followed by a dilation of A by B. The image processed by top-hat filter is the result the original image subtracts the original image's opening result image. Through this method, it can suppress the background of the image.

128.2.1.2 The Obtainment of ROIs

After suppressing the background, the method of adaptive threshold segmentation is adopted to get the ROIs. The binarization threshold can be obtained by using (128.2):

$$TH(k) = m_1(k) + t\sigma(k), \tag{128.2}$$

where $TH(k)$ is the binarization threshold of the kth frame, $m_1(k)$ is the mean value of the kth frame, and $\sigma(k)$ the standard deviation of the kth frame. t is a constant; it often takes 3–6. According to (128.2), it can get the threshold and segment the ROIs from the image. Results are shown in Fig. 128.2.

128.2.2 Enhancing the SNR to Get the Candidate Targets

After getting the ROIs, we need to process the ROIs according to their visual saliency. This article adopts the Dog filter which is the on-center filter and the off-surround filter in the scale space to process them. It not only can suppress the

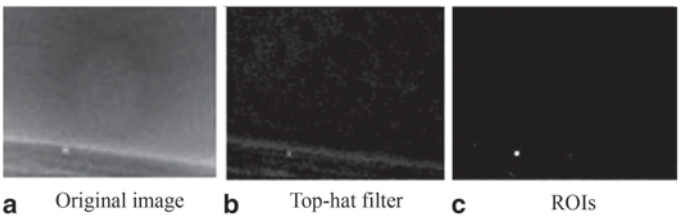


Fig. 128.2 The results of the ROIs. **a** Original image. **b** Top-hat filter. **c** ROIs

background but also enhance the targets. After getting the candidate targets in the scale space, the real targets can be selected by processing the SCRs of them.

128.2.2.1 Dog Filter in the Scale Space

It is necessary to get more image details when detecting the targets. The scale space theory brings a parameter viewed as scale in the image to build the model. By changing this parameter continually, we can get more image information.

Gaussian convolution kernel is the only linear kernel to achieve scale transformation, and the Dog filter is based on it. It is defined as (128.3)

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma), \quad (128.3)$$

where $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)} / 2\sigma^2$. It is the Gaussian function whose scales are variable; (x,y) is the spatial coordinate; σ is the standard deviation, which determines the smoothing degree of the image. Large scale means the general features of the image, and small scale means the details of the image.

When the center of the Dog filter and the center of the target are the same, a peak response will appear. According to this feature, it can detect the location of the targets. And it is found that there are some connections between the scale of the gauss and the scale of the target [13]. It has been shown as (128.4)

$$h = 2\sqrt{2}s. \quad (128.4)$$

In the actual process, by subtracting two images obtained through two continuous scales of Gaussian filter, one Dog filter image can be obtained. After comparing each points' Dog response with other 26 points' Dog response, which are the eight points in its 9×9 neighborhoods in the current scale image and two adjacent scale image's nine points in the same location, the points with the local maximum response are selected as the candidate targets. The process is shown in Fig. 128.3.

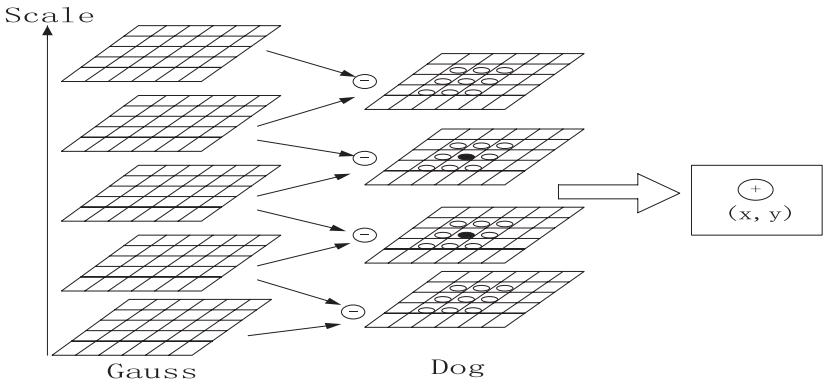


Fig. 128.3 The process of obtaining the candidate targets in the scale space

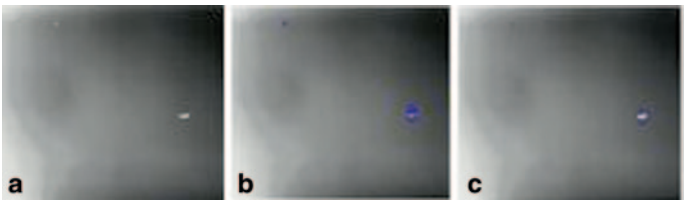


Fig. 128.4 Results of selecting the targets. **a** Original image. **b** Candidate targets by Dog. **c** Detection results

128.2.2.2 Selecting the Targets

After getting the candidate targets, there are also some noises which need to be excluded. By thresholding the targets' local SNR, the true targets can be identified. The local SNR of the target is calculated as (128.5)

$$SCR(x, y, \sigma) = \frac{|\mu_T - \mu_B|}{\sigma_B}, \tag{128.5}$$

where μ_T and μ_B are the mean value of the candidate targets and background, respectively. Each candidate target satisfies $|\mu_T - \mu_B| = D(x, y, \sigma)$, σ_B is the standard deviation of the background which can be calculated by the image in the scale space. The domain of the target and background are set to be $2\sqrt{2}\sigma$ and $4\sqrt{2}\sigma$ respectively. Afterwards, the points whose SCRs are greater than the threshold and which have local maximum SCRs are selected as the targets. The threshold value in this article is set to be 3. The results of the selection are shown in Fig. 128.4.

128.2.3 Accurate Detection of Using Retinex

Through the above process, most of the targets can be detected, but it is difficult to detect the targets almost submerged in the background. Considering this, this article proposes enhancing a certain area of the obtained target in the last frame using the theory of MSR and segmenting the targets from the local area.

The theory of Retinex is based on the theory of color constancy. By simulating human visual constancy, it can compensate the details of the image and enhance it. The principle of the single-scale Retinex theory can be expressed as (128.6)

$$TH(k) = m_1(k) + t\sigma(k), \tag{128.6}$$

where $S(x, y)$ is the input image distribution function. $F(x, y) = \lambda e \frac{-(x^2 + y^2)}{c^2}$, it satisfies $\iint F(x, y)dxdy = 1$. *r(x, y)* is the enhanced image we get.

The theory of MSR has the advantage of single-scale Retinex in low, middle, and high scales. It is defined as (128.7):

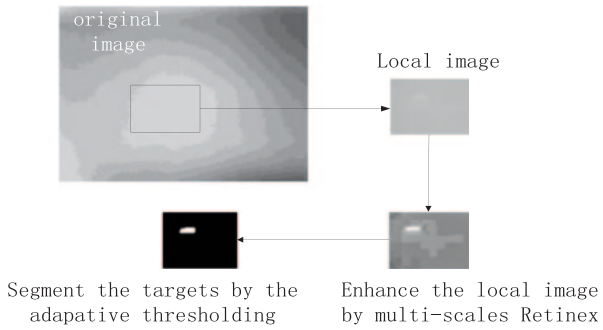


Fig. 128.5 The enhancement image by multi-scale Retinex

$$r(x,y)=\sum_{k}^{K}\omega_k\{\log S(x,y)-\log[F_k(x,y)*S(x,y)]\}. \tag{128.7}$$

Generally, k takes 3. $\omega_1=\omega_2=\omega_3=1/3$ It can have good effects if $C_k(k=1,2,3)$ in $F_k(x,y)$ takes 15, 80, and 120, respectively.

In the algorithm, the steps of MSR for the image are as follows:

a. Detecting the targets in the current frame by the method of 2.1 and 2.2. If there is no target and the targets in the last frame detected are not at the border of the image. It means that in this frame the target's SNR may be too small to detect, then go to step *c*. If there are targets in this frame, go to step *b* and continue to judge.
b. Traversing each target detected in the last frame, judge whether there is a target in this frame close to it. Only if the target in the last frame does not appear at the border of the image and there is no target close to it, go to step *c*.
c. Enhancing a local region surrounding the target in the last frame using the theory of MSR. The region can be set to be several times of the target's size; in this article the region is set to be five times of the target's size.
d. Segmenting the target in the local region using the method in 2.1.2.

Figure 128.5 shows the final results using the above-described method.

128.3 Experimental Results

In this section, there are three experiments to demonstrate the detection accuracy of the algorithm. In the first experiment, the algorithm is tested by a variety of background images. The results are shown in Fig. 128.6.

In the second experiment, four frames (66, 110, 377, 548) from the video images whose frame rate is 25 frames/s and has a resolution of 320×240 to test the algorithm are selected. The proposed algorithm is compared with Kim's and Shao's

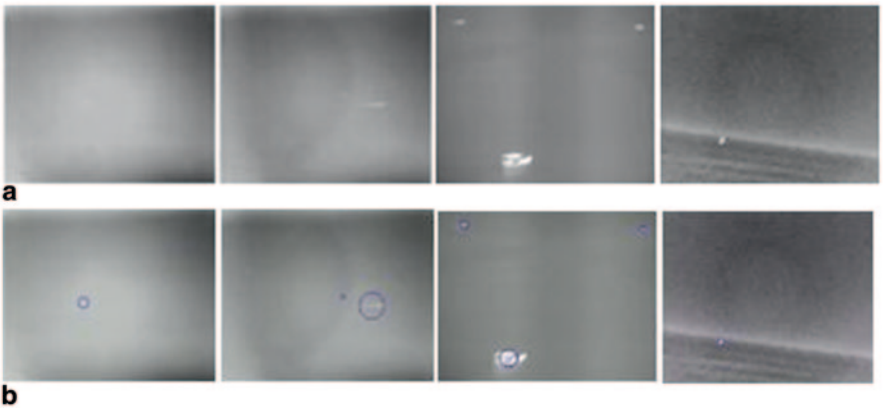


Fig. 128.6 Detection results. **a** Original images. **b** Detection results by the proposed algorithm

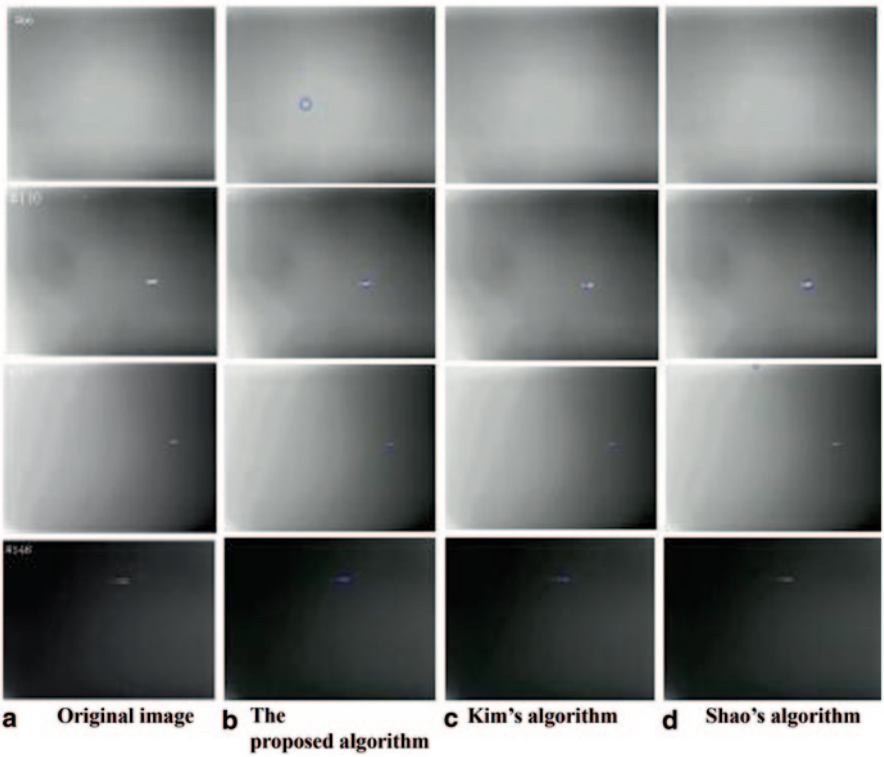


Fig. 128.7 Experimental results. **a** Original image. **b** The proposed algorithm. **c** Kim's algorith. **d** Shao's algorithm

algorithm. In Fig. 128.7, the targets are almost submerged in the background in frame 66, but it can only be detected by the proposed algorithm. In frame 110 and 548, Shao's algorithm will miss the targets. In frame 377, Shao's algorithm wrongly

Table 128.1 The FR and the FAR of the algorithm

Algorithm	Missing image numbers	Incorrectly detected image numbers	FR (%)	FAR (%)
Kim's	212	0	10	0
Shao's	217	19	10	0.9
The proposed algorithm	27	3	1.3	0.1

FR false negative rate, *FAR* false alarm rate

detects a target. Thus, when the target's and background's gray-level value are greatly different, the proposed algorithm and Kim's algorithm can effectively detect the targets, but Shao's algorithm will have sometimes false detections. When the target is almost submerged in the background, the proposed algorithm works best.

In the third experiment, in order to research the algorithm accurately, the indicators of the false negative rate (FR) and false alarm rate (FAR) to test the algorithm, which are calculated as (128.8), are used. The video images which involved 2121 frames are selected. Their frame rate is 25 frames/s, and they have a resolution of 320×240 to test, some of which have more than one target and some have no target. The results are shown in Table 128.1.

$$FR[\%] = \frac{\text{Number of missing targets}}{\text{Number of true targets}} \times 100.$$

$$FAR[\#/\text{image}] = \frac{\text{Number of incorrectly detected targets}}{\text{images}} \times 100. \quad (128.8)$$

In the table, we can see Shao's algorithm has a large FAR. The article's algorithm is similar to the Kim's algorithm in FAR. But the article's algorithm can greatly improve the FR.

128.4 Conclusion

Aiming at some infrared targets' SNR is low, this article proposes an algorithm of detecting the infrared targets based on the visual attention mechanism. By using the top-hat filter and the adaptive threshold segmentation method, the ROIs of the image can be segmented. After getting the ROIs of the image, the Dog filter in the scale space is used to improve the SNR of the image and detect the targets. In order to prevent the missing targets, the theory of MSR is used to research the missing targets. By comparison of the proposed algorithm with Shao's and Kim's algorithm, it demonstrates the effectiveness of the proposed algorithm.

Acknowledgments This research is sponsored by the Natural Science Foundation of China (No.61370180).

References

1. Kim S, Song T, Choi B. Spatio-temporal filter based small infrared target detection in highly cluttered sea background. 2011 IEEE 11th International Conference on Control, Automation and Systems Processing, IEEE Computer Society. p. 1142–46; 2011.
2. Yilmaz A, Shafique K, Shah M. Target tarcking in airborne forward looking infrared imagery. Image Vis Comput. 2003 21(7):623–35.
3. Ghaeminia MH, Shokouhi SB. Adaptive background model for moving objects based on PCA. 2010 6th Iranian, Machine Vision and Image Processing (MVIP).p. 1–4; 2010.
4. Bae T, Zhang F, Kweon I-S. Edge directional 2D LMS filter for infrared small target detection. Infrared Phys. Technol. 2012;55(1):137–45.
5. Reljin N, McDaniel S, Pokrajac D, et al. Small moving targets detection using outlier detection algorithms. SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, p. 769804; 2010
6. Liou RJ, Azimi-Sadjadi MR. Multiple target detection using modified high order correlations. Aerospace Electronic Syst. IEEE Transac on 1998;34(2): 553–68.
7. Nishiguchi K I, Kobayashi M, Ichikawa A. Small target detection from image sequences using recursive max filter. International Society for Optics and Photonics, SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation, p. 153–66; 1995.
8. Wang X, Lv G, Xu L. Infrared dim target detection based on visual attention. Infrared Phys Technol. 2013;55(6):513–21.
9. Duan H, Deng Y, Wang X, et al. Small and dim target detection via lateral inhibition filtering and artificial bee colony based selective visual attention. PloS ONE. 2013;8(8):e72035.
10. Yao X, Li D, Sun X. Detection of small target in infrared image sequences using attention mechanism. Systems and Control in Aerospace and Astronautics, 2006. ISSCAA 2006. 1st International Symposium on IEEE, p. 456–60; 2006.
11. Kim S, Yang Y, Lee J, et al. Small target detection utilizing robust methods of the human visual system for IRST. J Infrared, Millim Terahertz Waves. 2009;30(9):994–1011.
12. Shao X, Fan H, Lu G, et al. An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system. Infrared Phys Technol. 2012;55(5):403–8.
13. Chen G, Guo L, Han J, et al. Infrared dim small target detection based on morphological band-pass filtering and scale space theory. Acta Optica Sinica. 2012;32(10):144–51. (In Chinese)

Chapter 129
Reconnaissance Technology about Reflective Panoramic Vision for Unmanned Platform

Jian Wang, Zhenhai Zhang, Kejie Li, Haiyan Shao and Tao Xu

Abstract The environment perception technology is most important key technology in the application of unmanned platform, and the panoramic vision imaging technology is one of most important methods in the environment perception; but the hyperbolic and parabolic single viewpoint reflective panoramic vision system used most widely have some defects, such as requiring expensive imaging lens or high precision installation, etc. On this basis, we propose the multi-viewpoint parabolic reflective panoramic vision system without expensive telecentric lens and high precision installation, present parametric model of imaging system, and analyze properties of the caustic surfaces. In the end, we design the multi-viewpoint parabolic reflective panoramic vision system to achieve the relational expression of the parameters about the maximum horizontal detection range and resolution, and analyze the affecting facts and then do experimental verification. The results indicate that the targets in farther position can be detected by using this multi-viewpoint parabolic reflective panoramic vision system.

Keywords Unmanned platform · Caustic surface · Virtual projection center · Multi-viewpoint

129.1 Introduction

With the fast development of information technology, the unmanned platform is widely used in military and civilian fields, such as unmanned aerial vehicles and unmanned vehicles. The environmental perception is the most important key technology of unmanned platform. It can enhance mobility of unmanned platform and improve ability of monitoring and tracking to targets. The panoramic vision system

Z. Zhang (✉) · J. Wang · K. Li · H. Shao · T. Xu
School of Mechatronical Engineering, Beijing Institute of Technology, 100081 Beijing, China
e-mail: tigerzzh@126.com

J. Wang
e-mail: wj707115@126.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_129

plays an important role in the application of unmanned platform because of its large field of view of 360° [1].

The reflective panoramic vision system can be divided into two kinds: single viewpoint and multi-viewpoint based on the number of viewpoint. Currently, the hyperbolic and the parabolic single viewpoint reflective panoramic vision system are the main types with the former combining perspective lens with hyperboloid reflector while in conformity with the perspective imaging model. It is more useful because the perspective lens feature low price and is easy to buy; however, the optical center of perspective lens and the focus-center of hyperboloid mirror must be located at the two focal point of hyperboloid during installation; otherwise, the multi-viewpoint imaging will be caused; and the latter combines telecentric lens with paraboloid reflector. It conforms to parallel projection imaging model while having more compact structure, good image quality; but the telecentric lens is more expensive [2].

Because of the shortcomings of single viewpoint reflective panoramic vision system, this chapter designs the multi-viewpoint parabolic reflective panoramic vision system by means of combining lens with parabolic reflector. The system has both good imaging effect and economic practicality.

129.2 Multi-Viewpoint Parabolic Reflective Panoramic Vision System

The parabolic reflector is processed and polished by high precision numerical control machine, and its surface is plated with chromium by vacuum evaporation. As shown in Fig. 129.1, the charge coupled device (CCD) camera has been placed vertically on a fixed base, and optical axis of lens aligned with the centerline of

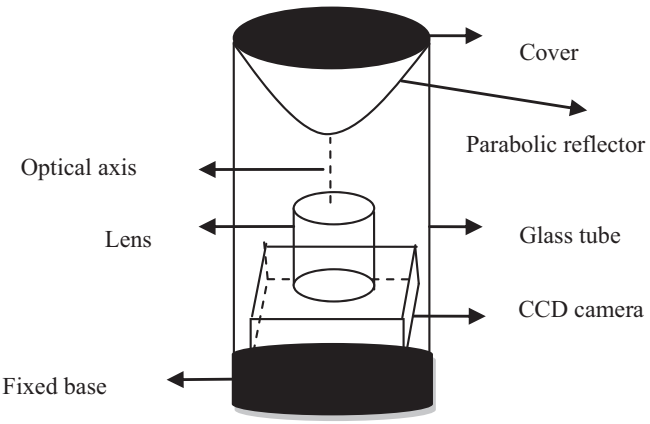


Fig. 129.1 Schematic diagram of multi-viewpoint parabolic reflective panoramic vision system. *CCD* charge coupled device

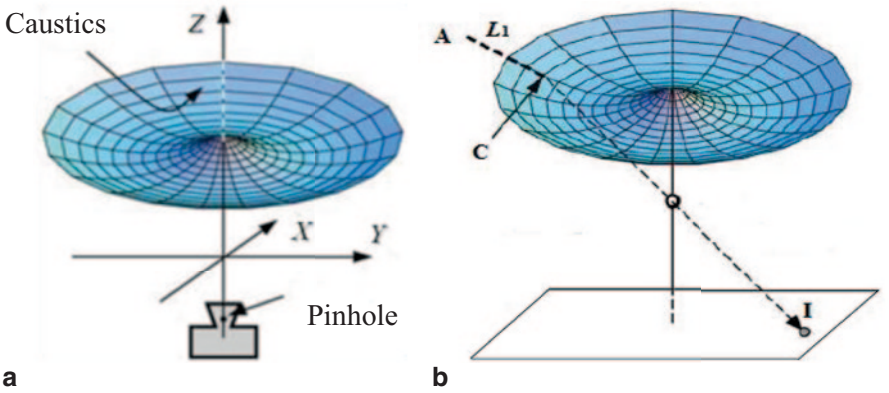


Fig. 129.2 Viewpoint distribution of the system

parabolic reflector. The reflection picture on reflector could be clearly imaged by adjusting the lens aperture and depth of field. Finally, with the reflection picture captured, the panoramic images with a wide visual field of 360° can be acquired continuously.

129.2.1 Analysis of Viewpoint Distribution

The multi-viewpoint parabolic reflective panoramic vision system does not have single viewpoint, Rahul [3–4] proved that tracks of all viewpoints will form curved surface known as caustics. Each point on the caustic surface represents the three-dimensional position of a viewpoint, as shown in Fig. 129.2a. According to Fig. 129.2b, point A in the scene along the incident ray L_1 is tangential to the caustic surface, and the tangential point C is the viewpoint of the incident ray.

129.2.2 Analysis of Imaging Model

Since caustics of system has rotationally symmetric characteristic, calculating and deriving can be done in the two-dimensional coordinate system xOz. From Fig. 129.3, world point A along the incident ray L_1 intersect parabola at A_2, is tangential to the caustic surface at C. Reflected ray L_2 entering optical centre F images on the CCD, the image point is I. To calculate the coordinate of point A, we can use point C as the center of projection to solve the single viewpoint reflective panoramic vision system.

Due to the reversibility of optical path, the coordinate of t point A corresponding to point I can be acquired by inverse projection. The incident ray L_1 intersect projective cylinder at A_1, A_1 is the projection point to be solved. Referring to Fig. 129.3,

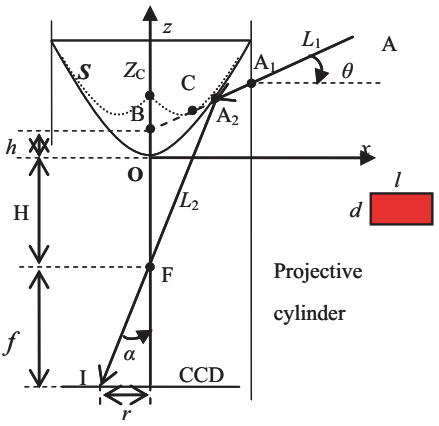


Fig. 129.3 Imaging principle diagram of system

the coordinate of point I and point A_2 are $(-r, -H + f)$ and (x_2, z_2), the distance between optical centre and the vertex of parabola is H, the focal length is f.

Let the parabolic equation be given by $x^2 = 2pz$. The gradient K_2 of reflected ray L_2 is:

$$k_2 = \cot(\alpha) = \frac{f}{r} = \frac{H + z_2}{x_2}. \tag{129.1}$$

So, the equation of reflected ray is:

$$z = \frac{H+z_2}{x_2} x - H. \tag{129.2}$$

From parabolic equation, we derive tangent gradient of parabola at point A_2:

$$k = \frac{dz}{dx} = \frac{1}{2p} \times 2x \Big|_{x=x_2} = \frac{x_2}{p}. \tag{129.3}$$

The normal line equation of parabola at point A_2 is then given by:

$$z = -\frac{p}{x_2} x + z_2 + p. \tag{129.4}$$

From Eq. (129.2) and (129.4), we solve for the gradient K_1 of incident ray L_1:

$$k_1 = \frac{x_2^4 + (3p^2 + 2pH)x_2^2 - 2p^3 H}{2p(p^2 + 2pH)x_2}. \tag{129.5}$$

The equation of incident ray L_1 is derived as:

$$z = k_1 x + (z_2 - k_1 x_2). \tag{129.6}$$

So, the coordinate of projection point A_1 can be got.

129.2.3 Properties of Caustic Surfaces

Rahul presents some characteristic properties of caustics such as surface singularities in his study. The singular point is only located on the optical axis in rotationally symmetric system. Thus, there is singular point on caustic surfaces which is called a cusp in the paper, as shown in Fig. 129.3. Let the coordinate of Z_C be (0, z_C), the extended line of A_1 A_2 intersect z axis at B, the coordinate of B is (0, h). From Eq. (129.6), we get:

$$h = z_2 - k_1 x_2. \tag{129.7}$$

We substitute parabolic equation and Eq. (129.5) into (129.7), to get h:

$$h = -\frac{2}{p + 2\mathrm{H}} z_2^2 - \frac{2p}{p + 2\mathrm{H}} z_2 + \frac{p\mathrm{H}}{p + 2\mathrm{H}}. \tag{129.8}$$

As seen from Eq (129.8), h is the quadratic function on z_2, and its maximum value is the z coordinate of the cusp point z_C, therefore, the coordinate of z_C is (0, pH/(p +2H)). In practical applications, the panoramic vision systems mostly have a projection center; however, the parabolic reflective panoramic vision system designed in this chapter does not have a single viewpoint, and does not have a unique projection center. In addition, the multi-view panoramic vision system will cause parallax. So we need to assume that cusp z_C as the virtual center of projection, which is similar to the single viewpoint projection model.

129.3 Analysis of Parameters

129.3.1 Maximum Horizontal Detection Range

Referring to Fig. 129.3, the red rectangle denotes the side face image of rectangular parallelepiped target, its area is $\delta = l \times d$ ($l \geq d$). From parabolic equation, Eqs. (129.1) and (129.5), we derive that the world points which vertical field angle is θ can be imaged on CCD, and image points form a circumference which radius is r. So there exists a one to one correspondence between θ and r [5]. In practice, we think that the target can be observed effectively, when the length and height of

target occupies at least two pixels on image plane respectively on CCD (pixel size is μ). In this case, the slant range U between target and the virtual center of projection is called the maximum horizontal detection range when the vertical field of view is θ. As U is much larger than the length, breadth, and height of target, the maximum horizontal detection is derived by the geometric proportions:

$$\frac{2\mu}{2\pi r} = \frac{d}{2\pi U \cos\theta}, \tan(\theta) = \frac{z_2 - h}{x_2}. \tag{129.9}$$

From Eq. (129.9), we solve for the U:

$$U = \frac{rd}{2\mu \cos\theta}. \tag{129.10}$$

Therefore, the premise is to guarantee clear panoramic image. The maximum horizontal detection of system is relevant to the size of the target itself, relative position between system and target.

129.3.2 Resolution

From Fig. 129.3, the red rectangle denotes target, the distance between perspective lens and target is given by U_1, which is much larger than the length and height of target. The angles between z-axis and reflection lights which are reflected by incident lights across top and bottom edge are defined α_2 and α_1, and they can be solved by Eqs. (129.1) and (129.8) [6]. Hence, the length i and height j of target on image plane can be calculated as:

$$j = f(\tan(\alpha_2) - \tan(\alpha_1)). \tag{129.11}$$

As the panoramic image has image distortion and it is closer from the center of target surface, the distortion is more serious, and the bottom edge of target image is closer than top edge; thus, we can calculate the length by using α_2, which is derived by the geometric proportions:

$$i = \frac{lf \tan(\alpha_2)}{U_1}. \tag{129.12}$$

So, the resolution is derived as:

$$\mathrm{R} = \frac{i \times j}{l \times d}. \tag{129.13}$$

Therefore, if the f is certain, the resolution is relevant to the size of target, relative position between system and target.

129.4 Experiment Results and Analysis

Figure 129.4 shows multi-viewpoint parabolic reflective panoramic vision system designed in this chapter, where $p = 15.75$ mm (the parameter of parabolic reflector). In order to use the maximization of CCD target surface, and not decrease strength of system, let $f = 12$ mm, $H = 31$ mm.

We select a moving car as target whose length l is 3.81 m and the height d is 1.86 m. Figure 129.5 shows a partial graph of the experimental results, target size is 3×2 pixels. The maximum horizontal detection distance is measured of 490 meters, $r = 630$ pixels is obtained by experimental results. Figure 129.6 shows the simulation results; therefore, the experimental results agree with the simulation results indicating that the targets in farther position can be detected.

Figure 129.7 shows partial graphs of panoramic image at 250 m, 300 m and 400 m, respectively; accordingly, the target size are 10×5 pixels, 8×4 pixels and 5×3 pixels. From Fig. 129.8, we can conclude that the resolution is relevant to the distance between system and target; and the resolution is descending as the distance increases.

Fig. 129.4 Multi-viewpoint parabolic reflective panoramic vision system

Fig. 129.5 Test graph of the maximum horizontal detection

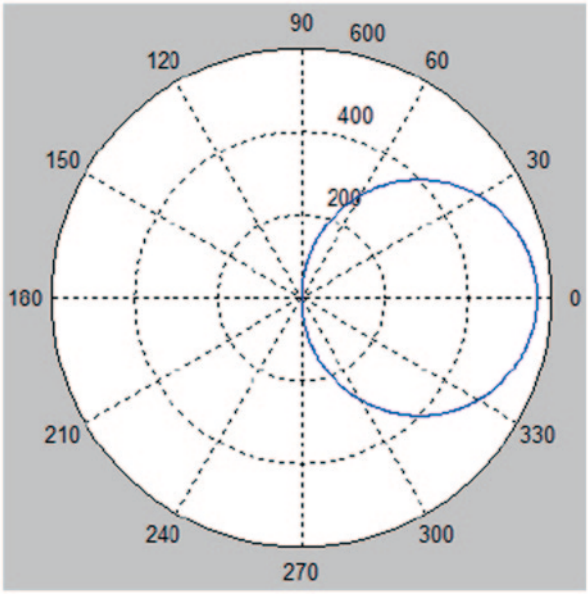


Fig. 129.6 Simulation results

Fig. 129.7 Partial graphs of panoramic image at 250 m, 300 m and 400 m

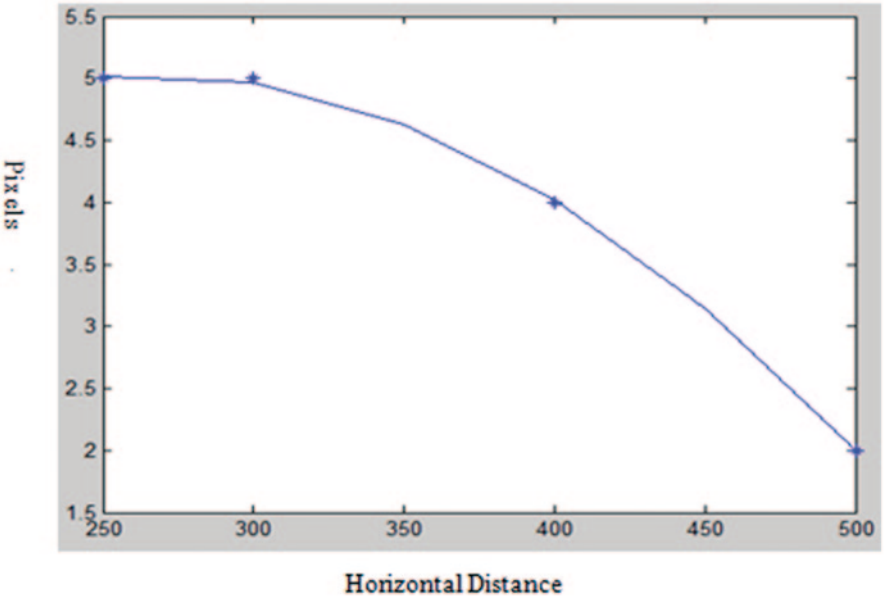


Fig. 129.8 Resolution vs. horizontal distance

129.5 Conclusion

The multi-viewpoint parabolic reflective panoramic vision system designed by us will avoid some defects of the hyperbolic and parabolic single viewpoint reflective of the panoramic vision system such as requiring expensive imaging lens or high precision installation, and the maximum horizontal detection range can reach 500 m approximately, so the targets in farther position can be detected by this system.

References

1. Zhang M, Liu PZ. The key technology research of environment perception on unmanned ground vehicle. Veh Power Technol. 2007;106(2):44–51. (In Chinese)
2. Xu W, Gao H, Wang W. Design and realization of multi-viewpoint catadioptric omnidirectional imaging system. J Syst Simul. 2010;22(2):435–8. (In Chinese)
3. Swaminathan R, Grossberg MD. Non-single viewpoint catadioptric cameras: geometry and analysis. Int J Comp Vis. 2006;66(3):211–29.
4. Swaminathan R, Grossberg MD. Caustics of catadioptric cameras. Computer vision, 2001. ICCV 2001. Proceedings Eighth IEEE International Conference on, IEEE. vol. 2, p. 2–9; 2001.
5. He SH, Yang SQ, Shao XF. A model of horizontal detection range for hyperbolic-catadioptric panorama image system. J Detect Control. 2011;33(3):39–42. (In Chinese)
6. Ma ZL, Wang JZ. Reconnaissance technology of catadioptyic panoramic vision for combat robot. Acta Armamentarii. 2011;32(4):503–7. (In Chinese)

Chapter 130
Platform Construction of Vegetable Safety Traceability

Haiting Diao, Chao Dong and Yimin Nie

Abstract In recent years, the food safety problems emerge in an endless stream. We urgently need to solve the problem of food safety to ease public trust crisis. This chapter discusses some key techniques of vegetable safety that relates to the whole platform. The whole system adopts a three-layer architecture: the data access layer, business logic layer, and user interface. According to the survey, in authority, we adopt the three levels of the general user, enterprise, and government. On this basis, we complete the vegetables space back, vegetable information query, data statistics, the vegetable transportation route analysis, and other functions. This chapter shows all links of information of vegetables from planting to the last. The platform of the assistant decision function is very beneficial to some extent for the government and the relevant government departments to supervise the vegetable safety.

Keywords ArcGIS Server · SilverLight · Quality and safety · Traceability · Platform

130.1 Introduction

Due to the speedy development of Chinese economy and the escalation of people's living standard, people pay more attention to the quality and safety of agricultural products. In recent years, a series of food safety incidents occurred in the "poison milk powder," "Clenbuterol," "sulfur ginger," "drainage oil," "dyeing bread," "beef paste," "medicine with apple," etc. Recently, the illegal use of copper sulfate for

H. Diao (✉) · C. Dong
College of Information Science and Academic, Shandong Agricultural University,
271018 Tai'an, China
e-mail: htdiao@sdau.edu.cn

C. Dong
e-mail: marchac-0968@163.com

Y. Nie
College of Resources and Environment, Shandong Agricultural University,
271018 Tai'an, China
e-mail: ymnie@sdau.edu.cn

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_130

industrial use was traced to shorten the curing cycle of preserved eggs in some egg processing enterprises. The frequent occurrence of food safety incidents has erupted public crisis of confidence that almost spook the people's consumption. It is urgent to solve the problem of food safety. Therefore, people also raised the research boom of food safety in china. In fact, as early as the 1990s, many countries have carried out the research of food traceability system, such as the European Union, the USA, Canada, Japan, Australia, Korea, and so on. In China, the quality and safety of agricultural products traceability started late; there are many imperfections compared with developed countries. But the government and scholars have done a lot of research and implementation[1].

At present, the establishment of a traceability system is one of the effective ways to solve the problem of food safety. So, how to establish an effective system is worth to discuss. Some people built a system which mostly focused on the access to product traceability information. These systems were mainly embodied as tabular data, a few combined with spatial data. Even if some people in the construction process used the geographic information system (GIS) technology, they just displayed the map with a lack of correlation analysis with GIS technology. This chapter focuses on the key technology, the function, some problems, and solutions of vegetable traceability platform, especially the application of Silverlight and WebGIS.

130.2 Some Key Technologies of the Platform

130.2.1 ArcGIS Server

The ArcGIS Server is an enterprise GIS product on the Web by the Environmental Systems Research Institute (ESRI) company, which can be used for GIS applications' network publishing and processing. ArcGIS Server builds centralized management and supports for multiple users. It provides an efficient framework platform for creating and managing GIS application based on the server. The ArcGIS Server is a distributed system, which is distributed in various roles of multiple machines for collaborative work. The function of GIS is mainly reflected in the server implementation and management. At the same time, the ArcGIS Server is a server manager, to manage the various geographical resources.

Developers use the ArcGIS Server to build Web applications, Web services, and other enterprise application on .NET and J2EE on the Web server. The ArcGIS Server fully embodies the advantages of ESRI software architecture based on the standard. It makes full use of ArcObjects—the core components library of the ArcGIS. The ArcGIS Server provides WebGIS services based on industry standards. We can use a variety of development environments, application servers, and database management systems (DBMSs) to achieve it [2].

130.2.2 Cloud Computing

Cloud computing technology is very popular in recent years. Sometimes cloud computing is compared to the third wave of information technology, after the personal computer and the Internet. That remarkably shows its important status in the information industry. Since the "cloud computing" concept is proposed, it attracted more and more attentions of the government and community. In many countries, companies have done a lot of research. Reactions to the concept of cloud computing vary widely. The definition from the National Institute of Standards and Technology is as follows: Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This concept points out the essence of cloud computing that we put information technology (IT) resources as a service to provide on the Internet. Under the cloud computing environment, everything is a service [3].

130.2.3 ArcGIS API for Silverlight

ArcGIS application program interface (API) for Silverlight is produced by the ESRI Company, which is used to develop the WebGIS application on the Silverlight platform for a set of programming interfaces. ArcGIS API for Silverlight visits the ArcGIS Server map service, video services, geometric services, geoprocessing services, feature services, network services through the representational state transfer (REST) interface. It can realize query and retrieval, network analysis, temporal perception, and other functions [4].

130.3 System Framework Design

The whole framework breaks through the restriction of the two-layer model, which consists of the client/server (C/S) and the browser/server model. It adopts a three-tier architecture, which is the data access layer (DAL), business logic layer (BLL), and the user interface (UI) from bottom to top. This three-layer architecture is beneficial for the system development, maintenance, deployment, and expansion, which makes it easy to allocate resources and reflect the low coupling characteristics.

The bottom of the system is the data access layer, which is the manipulation layer for various data. It includes the storage, management of the spatial data, and the attribute data, while providing data services for the business logic layer. The attributed data is stored in the database. As far as possible, we separate the GIS and conventional Web development. We used the ArcGIS online maps.

The middle layer is the business logic layer. It creates the system and data separation. It builds a bridge between the data access layer and the user interface. It avoids direct access to data that causes data loss or damage. It also lays a good foundation easy to expand and to maintain the system.

The top layer is the user interface, which means that the client part is realized in the application. That also refers to our receptionists. It is responsible for direct interaction with the user, management, and the Web page. It sends the user input to the business logic layer, to receive user input data, and displays the processed data desired by users. In this case, we use controls of ESRI Silverlight API to achieve the map interaction between the users (Fig. 130.1).

130.4 Function

130.4.1 System Function Design

According to the authority, we divided the ordinary users, production enterprise users, and government regulation users.

For ordinary users with minimal permissions, you can browse according to the trace code to get the vegetable basic information and vegetables logistics information.

For production enterprises users with some senior permission, you can also upload product information in addition to the basic query.

For government regulators with the highest authority, you belong to the administrator user. You can maintain the user information and publish the vegetable safety analysis.

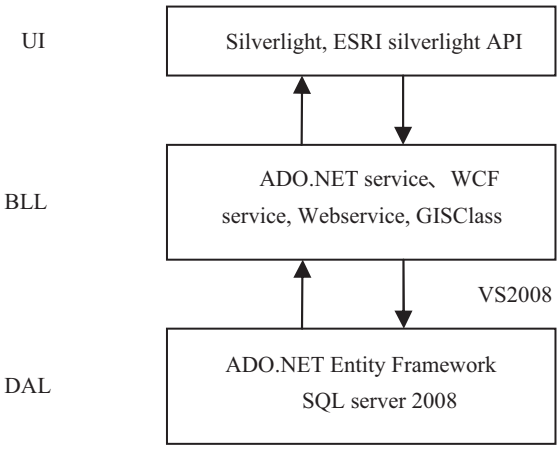


Fig. 130.1 The overall frame structure chart of the system. *UI* user interface, *BLL* business logic layer, *DAL* data access layer, *ESRI* Environmental Systems Research Institute, *API* application program interface, *GIS* geographic information system, *SQL* sequential query language, *WCF* Windows communication foundation

130.4.2 Major Functions

130.4.2.1 Information Input Function

We take the vegetables information into the database through this module, for later query or analysis.

130.4.2.2 Product Query

The query is the most basic function of GIS. The system can be queried according to the bar code or name, showing the query results in the form of a table.

130.4.2.3 Space Traceability

According to the bar code, we can trace back, and we connect the logistics node, then check the logistics distribution path. In the development process, we use the ToutTask class of ArcGIS API for Silverlight. The specific operation is that the users input the bar code information and click the check button; then information is displayed in the table. Then you click the button of the show path, the map will show the way to link the path. Then a car icon will travel according to the path. The speed of the car can be controlled. Compared with the traditional form, the display mode and map are combined, with positioning information. Putting a car icon enhances the dynamic effect, which makes the back more intuitive (Fig. 130.2).

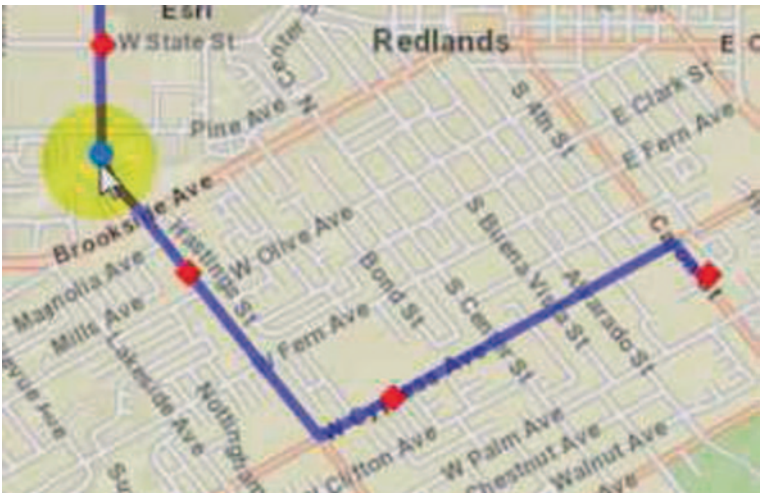


Fig. 130.2 Space traceability chart

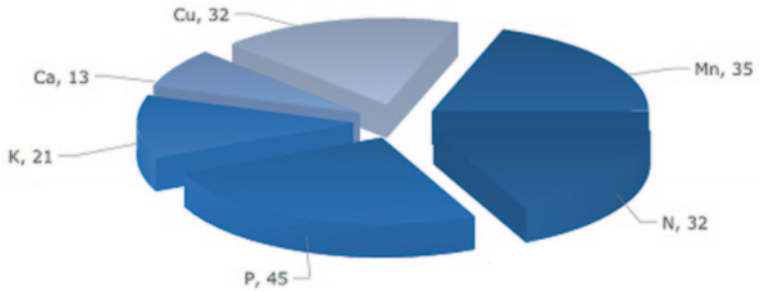


Fig. 130.3 Pie chart

130.4.2.4 Data Statistics

The system dynamically generates statistical charts about vegetables. Statistical charts can be a pie chart or a scatter plot, which makes data expression more vivid (Fig. 130.3).

130.4.2.5 Path Analysis

In this module, we can carry out the optimization design of the path according to the start and end points of vegetable transportation, such as the least time or the shortest distance. This can help to make transport personnel route decisions (Fig. 130.4).

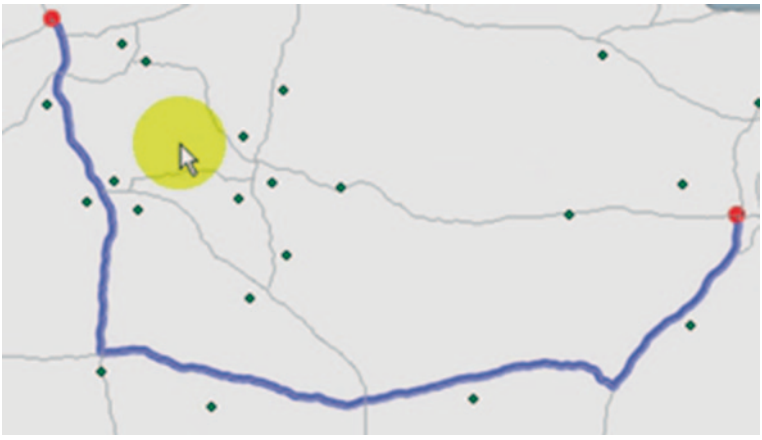


Fig. 130.4 The best path analysis diagram

130.5 Discussions

130.5.1 Making Full Use of the Spatial Analysis Function of GIS

The powerful spatial analysis function is the greatest advantage of GIS, compared with others. Therefore, how we can use GIS better for the vegetable safety service is a problem that is worth discussing. The platform makes full use of the spatial analysis function of ArcGIS Server. We developed a vegetable space back function, best path analysis function, etc. These functions provide a great technical support for decision making for the relevant departments.

130.5.2 The Stereo Back

Most systems have only the three-dimensional display of the vegetable origin. The platform presents different nodes of information data three-dimensionally that are integrated of space and time. The trace function of the vegetable was improved with the help of GIS, and truly realizes the three-dimensional information display.

130.5.3 Study the Combination of Cloud Computing and GIS

The goal of cloud computing is that through the network the user can maximize the use of virtual resource pools at any time, any place, processing large computational problems [5].

ArcGIS Online is a public resource website online in the ESRI Company, which provides a one-stop map of cloud services. This study related to the cloud computing is the use of ArcGIS Online and ArcGIS Web Mapping API. In the process of platform development, we visit the free high-quality map through the ArcGIS Online. At the same time, we combine two-dimensional vector road data for query and traceability, which enhances the query visual effect.

130.5.4 Data Interaction Problem

For security reasons, Silverlight is not allowed to directly access the database. Silverlight does not support Dataset, that is to say Silverlight does not fully support ADO.NET. Taking safety and efficiency under consideration, the Windows communication foundation (WCF) rich Internet application (RIA) Service is proposed. Containing the WCF component in Silverlight plug-in, developers can use

the requested data into a WCF service, to operate the database through the WCF service. It is said that the developers can create a WCF service operation of the database, and the Silverlight application program is to call these operations.

WCF RIA Service solves the problems of application logic layer and data layer interaction. It provides a safe and efficient solution for application access and modifies the database.

130.6 Conclusion

Vegetable quality and safety problems are related to the benefit for the people's livelihood. From the farm to table, vegetables go through a lot of links. Every link has the possibility of contamination and is very important, so we must complete all in one service. This system can provide consumers with vegetable product information query, provide the interface for the production management of manufacturing enterprise, and provide a good platform for the implementation of government regulation. A prominent characteristic in the process of system construction is that many functions are combined with each other and in the map, which makes up a single previous traceability system shortcoming. That makes the query and traceability more intuitive, so the user has better experience.

References

1. Zhang H, Sun X, Liu Y. Research on the feedback system for quality and safety of agricultural products [J]. Wuhan, Hubei Africultural Sci. 2010;49(12):3220–3 (In Chinese).
2. ESRI China (Beijing) Training Center. ArcGIS Easy Tutorial-ArcGIS Server. 2008. http://wenku.baidu.com/view/409d42d7240c844769eaeef8.html.
3. Liu R. Cloud computing and its key techniques [J]. Chengdu, Digitization User. 2013;40(6):15–16 (In Chinese).
4. ESRI China (Beijing) Limited Company. ArcGIS API for Silverlight Application Development Tutorial. 2011. http://wenku.baidu.com/view/d5de44db76eeaeaad1f3302e.html.
5. Li Q, Zheng X. Research survey of cloud computing [J]. Chongqing, Computer Sci. 38(4):32–37(In Chinese).

Chapter 131
A Mobile-Gateway-Assisted Transmission Scheme for Wireless Sensor Network

Lianhai Shan, Weidong Fang, Xiaoqing Mei, Yong Xiong and Weiwei Gao

Abstract Wireless sensor networks (WSNs) have been applied in many different areas. Data transmission quality of service (QoS) has become a hot research point for the WSN, especially for the emergency data. Many researchers focused on developing WSN access and routing algorithms for the transmission of QoS, but less research has been focused on the mobile user equipment (UE) acting as a gateway for WSN data transmission. In this chapter, we propose a UE-gateway-assisted adaptive transmission algorithm for the WSN, which designs a new signaling to select the UE-gateway as the access point. Simulation results show that the WSN system energy cost and data transmission delay can achieve a better system performance by using the adaptive scheme compared to the traditional scheme.

Keywords WSN · Mobile gateway · Base station · Energy consumption · Transmission delay

131.1 Introduction

With the development requirement of Internet of things (IOT), the wireless sensor network (WSN) has caused an increasing interest for industrial and research. The WSN has played a vital role in our daily lives, i.e., e-health care, environment monitoring, industrial metering, surveillance systems, etc. The WSN can be described as a network of sensing nodes and may enable interactions among persons, computers, and the surrounding environment [1]. Since the fundamental advantage of the WSN can be deployed in ad hoc manner, one great challenge is to create an organizational structure among WSN sensor nodes to guarantee the data transmission [2].

W. Fang (✉) · Y. Xiong
Shanghai Institute of Microsystem and Information Technology,
CAS, 200050 Shanghai, China
e-mail: wd_fang@126.com

L. Shan · X. Mei · W. Gao
Shanghai IoT Co., LTD., Shanghai Research Center for Wireless Communications,
200335 Shanghai, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_131

The existing research about the quality of service (QoS) of WSNs includes two categories: medium access control (MAC) algorithm improvement and routing algorithm improvement. For WSNs, MAC plays an important role in the communication of QoS, i.e., channel access, transmission delay, and energy consumption. The existing MAC algorithms for WSNs mainly include sensor-MAC (S-MAC), timeout-MAC (T-MAC), and Berkeley-medium access control (B-MAC) which cannot provide a QoS guarantee. Distributed and mobility-adaptive clustering (DMAC) is an efficient algorithm to reduce the energy consumption and transmission delay, which is designed for WSN data gathering [3]. Zebra-medium access control (Z-MAC) is a hybrid MAC protocol, which can switch between carrier sense multiple access (CSMA) and time division multiple access (TDMA) mode for the WSN node transmission access [4]. Current research of routing algorithm for QoS improvement mainly includes sequential assignment routing (SAR) [5], SPEED (a stateless protocol for real-time communication) [6], etc. SAR is the first routing algorithm to guarantee the WSN transmission QoS. It considers the energy, service type, and packet priority of WSN nodes. SPEED routing algorithm is proposed with providing end-to-end guarantee for real-time service [7]. This algorithm requires each node to maintain its neighbors and geographic information in order to find the efficient transmission path. Recently, another WSN QoS-aware algorithm is proposed, where real-time data traffic is mainly focused on the image or video sensors [8]. The algorithm of finding the efficient transmission path is proposed which can reduce the end-to-end delay for WSN data transmission [9]. And another algorithm provides a QoS-optimizing scheme by improving the sensor scheduling and data routing algorithm [10].

In all the above algorithms, the sensor data is transmitted to the WSN sink unit by multi-hop. The end-to-end transmission delay may become large, especially for the specific WSN areas based on belt-shaped topology or linear topology. Simultaneously, sensor nodes near the sink unit would forward data more than the nodes far from the sink node, and their energy will be exhausted more quickly, which will cause a disconnection problem for the WSN.

This chapter proposes an adaptive access scheme by using the mobile UE-gateway assisting the WSN data to reduce the transmission delay and reduce the WSN energy cost. Especially, when the sensor node has emergency data to transmit, the UE-gateway will transmit the emergency data to the server directly instead of using the conventional multi-hop transmission mode to the sink unit.

The remainder of this chapter is organized as follows: Sect. 131.2 describes the UE-gateway-assisted adaptive transmission scheme in details. Section 131.3 is the numerical results and performance analysis part. Section 131.4 concludes the chapter.

131.2 Proposed Scheme

131.2.1 System Scenario

The scenario of mobile-UE-gateway-assisted adaptive transmission scheme is shown in Fig. 131.1. The UE-gateway is in the WSN area and has a WSN and cellu-

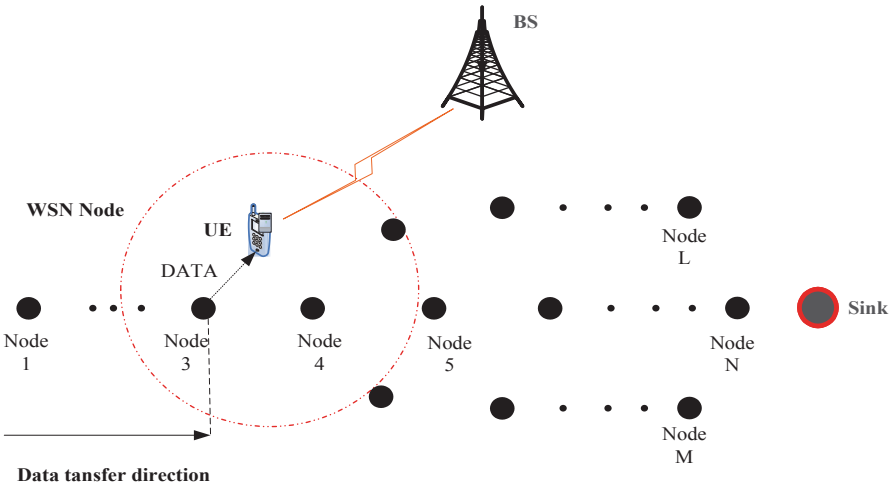


Fig. 131.1 User equipment (*UE*)-gateway-assisted transmission scenario. *WSN* wireless sensor network, *BS* base station

lar network interface, which can overhear the DATA frame from node 3 transmitting to node 4 in the WSN channel.

Afterward, it receives an acknowledgment (ACK) frame from node 4 transmitting to node 3. And then, if UE finds the data and emergent data type, the UE broadcasts transmission end-up signaling immediately and instructs node 4 not to receive the DATA frames from node 3. Finally, UE forwards all the DATA frames to the cellular base station (BS) by overhearing/receiving from node 3.

131.2.2 Adaptive Transmission Scheme Design

In the proposed algorithm, the UE-gateway does not need to execute the carrier sense multiple access-collision avoidance (CSMA-CA) and is able to use the channel preemptively, which will access the channel with the highest priority without any back-off period (non-CSMA). The UE-gateway broadcasts immediately the transmission end-up signaling on this channel after receiving the ACK frame from node 4 after it identifies that this data is an emergent data type, and the next hop sensor nodes cannot contend the channel successfully for data transmission. The detailed UE-gateway-assisted packet adaptive transmission process is shown in Fig. 131.2.

The UE-gateway-assisted adaptive transmission process is as follows:

Step 1 UE overhears a DATA from node 3 transmitting to node 4, and the UE overhears a corresponding ACK frame from node 4 transmitting to node 3.

Step 2 UE broadcasts end-up signaling and makes node 4 not to receive any frames except the beacon frames. In this state, node 4 still discards all data frames received

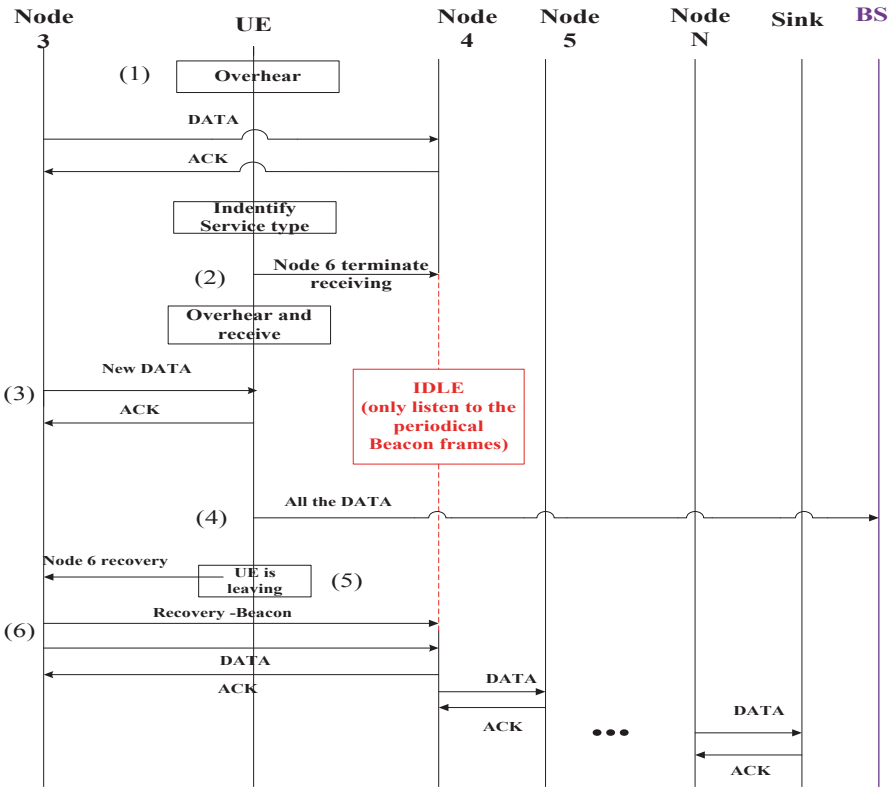


Fig. 131.2 User equipment (*UE*)-gateway-assisted transmission process. *ACK* acknowledgment, *BS* base station

from node 3. Node 4 can still wake up periodically according to the beacon interval of node 3. And it is not influenced if node 4 itself has data to transmit to the next hop.

Step 3 UE overhears the DATA frames from node 3 and responses an ACK frame to node 3 for each DATA frame. Simultaneously, node 4 does not receive any DATA except the beacon frames.

Step 4 UE forwards all the received DATA of node 3 to cellular BS directly, and then BS transmits them to the server.

Step 5 When the UE-gateway is leaving this area, it sends the leaving signaling to node 3 and makes node 3 send recovery signaling to node 4 because node 4 and node 3 are in synchronization by using the periodical beacon frame.

Step 6 Node 3 sends the recovery-beacon to node 4, and node 4 returns to normal working state. When node 3 sends DATA frames again, node 4 receives and transmits the data to its next hop. The sink nodes receive the data and then transmit them to the server for further processing.

131.3 Performance Evaluation

131.3.1 Simulation Scenario Description

The performance of the UE-gateway-assisted adaptive transmission scheme for the WSN is evaluated in this section. Two sensor networks of 100 and 200 sensor nodes are considered in simulation. In these cases, the nodes are distributed in a 200×200 m^2 square field. The initial energy of sensor nodes is 0.5 J. The chosen UE-gateway has enough energy.

In the WSN, a sensor node consumes $E_{Rx}(\mathrm{k}) = \xi_{elec} k$ J of energy for receiving a k-bit packet. At the same time, the energy consumption is $E_{Tx}(\mathrm{k},\mathrm{d}) = \xi_{\mathrm{elec}} k + \xi_{\mathrm{amp}} k d^2$ to transmit a k-bit packet to another node over a distance of d meters [11, 12]. In our simulation, we set $\xi_{\mathrm{elec}} = 50^{\circ}$ nJ/bit, and $\xi_{\mathrm{amp}} = 100^{\circ}$ pJ/bit/m^2. The simulation parameters are shown in Table 131.1.

The sensor nodes distribution scenario is shown in Fig. 131.3a and 131.3b, where the numbers of system sensor nodes are equals of 100 and 200.

Table 131.1 Simulation parameters

Parameter	Value
Field size	200 m × 200 m
Initial energy of nodes	0.5 J
Number of nodes	100 nodes, 200 nodes
Data packet size	Random (20, 100) Byte

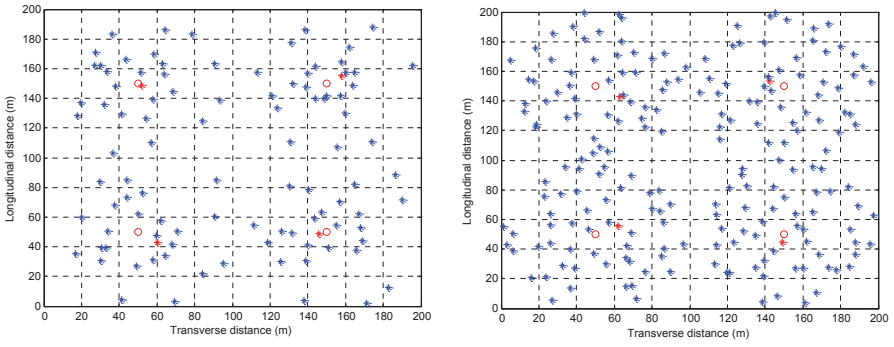


Fig. 131.3 WSN nodes distribution graph. **a** Number of system sensor nodes equals 100. **b** Number of system sensor nodes equals 200

131.3.2 Numerical Results

The sensor node transmits its packet to the UE-gateway if the gateway finishes the adaptive transmission process. Otherwise, it transmits the packet to the next hop according to the routing table of the WSN. The system energy cost with simulation time (rounds) can be described as follows: The energy cost of each transmission is mainly based on the transmission distance; the energy cost is exponentially incremental about the distance. In addition, the transmission hops are another important factor, and the energy cost increases with extra hops. From Fig. 131.4a and 131.4b we can see that the proposed algorithm can reduce the system energy cost because we use the UE-gateway to forward the detected data.

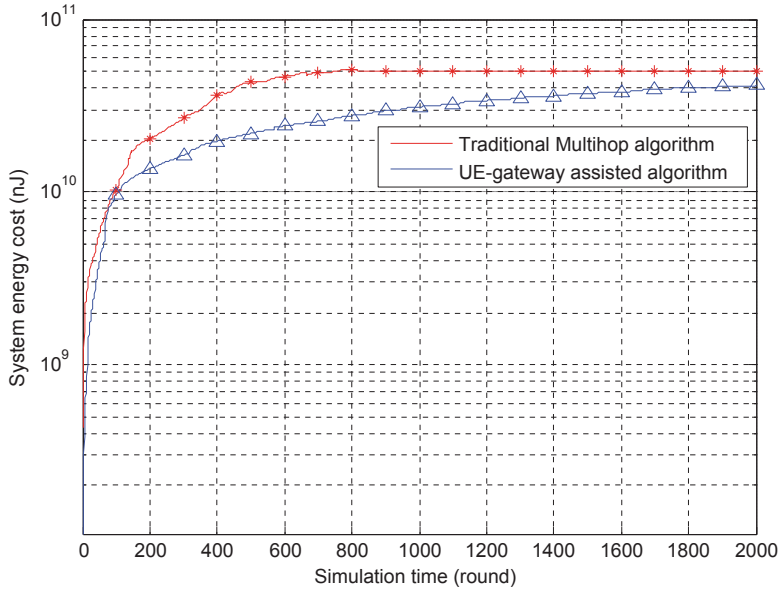
The main simulation results of different WSN sizes are shown in Table 131.2 and Table 131.3. The average energy consumption and transmission delay have the same trendline.

Since there are several UE-gateways in the WSN, the average delay of the WSN data transmission is smaller than the conventional WSNs with no UE-gateway, and the delay is reduced with the number of UE-gateway increasing. For the 200 nodes WSN, the average transmission distances between the nodes are shorter; the energy consumption per round increases about 75–90 % compared with 100-node network because it used more hops to forward the data. Moreover, the impact by using the UE-gateway for adaptive transmission that can be brought is obviously improved in energy consumption and transmission delay.

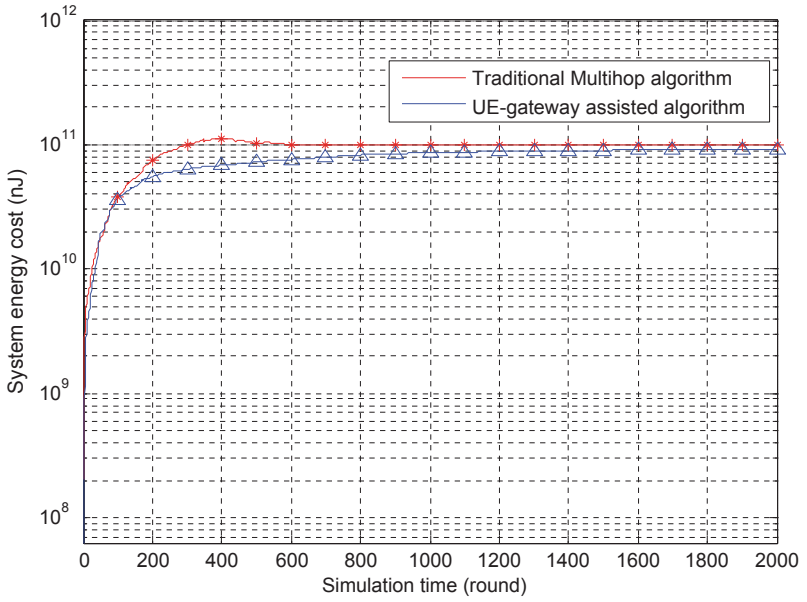
131.4 Conclusion

In this chapter, we proposed an UE-gateway-assisted adaptive transmission scheme for the WSN. The adaptive transmission scheme can reduce energy consumption and transmission delay largely for the WSN with UE-gateway assisted because the UE-gateway will forward the data to the BS directly. This scheme can be used in transmitting any WSN packet, and the UE-gateway will play a more and more important role for the IOT application.

Acknowledgments This work was supported by the National Natural Science Foundation of China (No.61302113), Shanghai Science Foundation (No. 13ZR1440800), Shanghai Rising-Star Program (14QB1404400), and Shanghai Science and Technology Innovation Program (No.12511503300, 12DZ2250200, 14511101303).



a Number of system sensor nodes equals 100



b Number of system sensor nodes equals 200

Fig. 131.4 System energy cost in the WSN. **a** Number of system sensor nodes equals 100. **b** Number of system sensor nodes equals 200. *UE* user equipment

Table 131.2 Performance comparison of 100 sensor nodes

Performance metrics		WSN with no UE-gateway	WSN with 4 UE-gateways
Network lifetime (rounds)	FND	73.8	244.9
	HND	388.3	952.0
	LND	958.6	2955.6
Energy consumption *(10^{-3} J/round)*		57.6	53.9
Delay *(slots/round)*		51.5	21.4

WSN wireless sensor network, *UE* user equipment, *FND* first node dies, *HND* half node dies, *LND* last node dies

Table 131.3 Performance comparison of 200 sensor nodes

Performance metrics		WSN with no UE-gateway	WSN with 4 UE-gateways
Network lifetime (rounds)	FND	43.7	82.4
	HND	240.5	418.4
	LND	582.5	1398.5
Energy consumption *(10^{-3} J/round)*		97.9	96.3
Delay *(slots/round)*		103.5	50.3

WSN wireless sensor network, *UE* user equipment, *FND* first node dies, *HND* half node dies, *LND* last node dies

References

1. Verdone R, Dardari D, Mazzini G, et al. Wireless sensor and actuator networks. 2nd ed. London: Academic; 2010. p. 50–65.
2. Ye W, Heidemann J, Estrin D. Medium access control with coordinated adaptive sleeping for wireless sensor networks. IEEE/ACM Trans Netw. 2004;12(3):493–506.
3. Lu G, Krishnamachari B, Raghavendra CS. An adaptive energy efficient and low-latency mac for data gathering in wireless sensor networks. In: Proc. Int. Parallel Distrib. Process Symp., IEEE, April 2004, Santa Fe, USA; 2004. p. 224–31.
4. Rhee I, Warrier A, Aia M, Min J. Z-MAC: a hybrid mac for wireless sensor networks. In: Proc. Sensys., ACM, November 2005, St Diego, USA; 2005. p. 90–101.
5. Zara H, Hussain FB. QoS in wireless multimedia sensor networks: a layered and cross-layered approach. Wirel Pers Commun. 2014;75(1):729–57.
6. Alam MM, Mahbub M. Energy-aware QoS provisioning for wireless sensor networks: analysis and protocol. J Commun Netw. 2009;11(4):390–405.
7. He T, Stankovic JA, Lu CY. SPEED: A stateless protocol for real-time communication in sensor networks. In: Proc. of the International conference on Distributed Computing Systems, IEEE, May 2003, Providence, USA; 2003. p. 46–55.
8. Mahapatra A, Anand K, Agrawal DP. QoS and energy aware routing for real-time traffic in wireless sensor networks. Comput Commun. 2008;29(4):437–45.
9. Singh J, Singh B, Chaudhary A. Ubiquity of mobile computing in wireless networks. Int J Technol Eng Syst. 1(1):1–4.
10. Zhu J, Papavassiliou S. Adaptive localized QoS-constrained data aggregation and processing in distributed sensor networks. IEEE Trans Parallel Distrib Syst. 2006;17(9):923–33.
11. Bouckaert S, Poorter ED. Strategies and challenges for interconnecting wireless mesh and wireless sensor networks. Wirel Personal Commun. 2010;53(3):443–63.
12. Gatzianas M, Georgiadis L. A distributed algorithm for maximum lifetime routing in sensor networks with mobile sink. IEEE Trans Wirel Comm. 2008;7(3):984–94.

Chapter 132
Adaptive Mobile Gateway: QoS-Guaranteed Challenges for Wireless Sensor Networks

Lianhai Shan, Weidong Fang, Wei Yao, Yong Xiong and Weiwei Gao

Abstract Quality of service (QoS) of wireless sensor networks (WSNs) communications has become one research hotspot, especially for the emergency data transmission. Many researchers focused on developing WSN routing algorithms to guarantee the transmission QoS, but less research has been concerned in the mobile user equipment (UE) acting as a gateway for WSN data transmission. In this paper, we first propose that the system application requirement for WSN, where UEs of mobile cellular network (MCN) act as sensor nodes and gateways for WSN. And then, we investigate QoS key technical challenges in the convergent networks of MCN and WSN. Finally, based on an analysis of technical challenges, we discuss the research points of QoS for WSNs.

Keywords WSN · Mobile UE-gateway · Base station · QoS · Machine-to-machine

132.1 Introduction

With the continued requirements of the machine-to-machine (M2M) development, Quality of service (QoS) of WSN has caused an increasing interest for industrial and research. In order to satisfy fast-increasing M2M communications requirements, multiple wireless networks convergence has become the trend [1, 2]. The QoS of WSN has also enormously promoted the development of M2M communications [3]. Generally, M2M is based on commonly and ubiquitously used technologies – wireless sensor network (WSN) and mobile cellular network (MCN), which usually uses the cellular system as the backhaul networks [4].

WSN is described as a network of sensing nodes and can interact with the surrounding monitored things [5]. Since the fundamental advantage of WSN can be

W. Fang (✉) · Y. Xiong
Shanghai Institute of Microsystem and Information Technology, CAS, 20050 Shanghai, China
e-mail: wd_fang@126.com

L. Shan · W. Yao · W. Gao
Shanghai IoT Co., LTD., Shanghai Research Center for Wireless Communications, 200335 Shanghai, China

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_132

deployed in ad-hoc manner, one great challenge is to create an organizational structure amongst WSN sensor nodes to guarantee the data transmission QoS [6]. For MCN, the profit brought by the WSN is mainly in applications extension. Comparably, with the help of MCN, the performance of WSN can be improved. The whole QoS architecture is shown in Fig. 132.1, and the converged networks make QoS more challenging to improve the WSN efficiency [7, 8].

132.2 QoS Requirements for WSN

In the new QoS-Guaranteed scheme for WSN, it includes two types of gateways: fixed gateway and mobile gateway, which is shown in Fig. 132.2. The gateways are dual-mode and have two interfaces: WSN and cellular networks [9]. Then, the data of WSN can be forwarded to the BS by the gateway. The mobile UE can also acquire the necessary information with the downlink data from the BS.

The WSN applications and services requirements are based on the existed WSN infrastructure and the perspective of the market they serve [10–13], which is shown in Table 132.1.

WSN is composed of lots of sensor nodes distributed in certain application area. Each sensor node has the capability of detecting data, such as temperature, pressure, humidity, noise etc. WSN nodes can send its detected data to sink node. Since most of the applications types have the same requirements on the network, we divide the existed data delivery models into four types: periodical, event-driven, query-driven, continuous mode [14, 15]. QoS requirements in WSNs mainly focus on the transmission delay and packet loss. Different applications have different QoS requirements, which we have expressed in the above section part. Since WSNs interact with the different environment [16, 17], the whole WSNs bring QoS challenges from different applications. Their particular characteristics are described as follows:

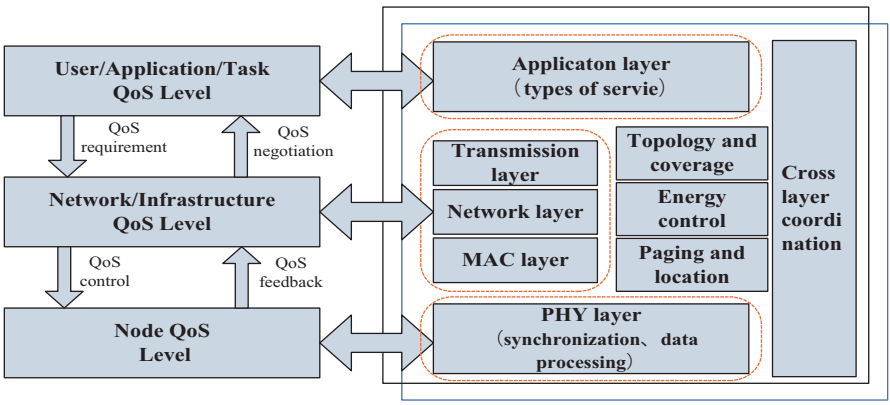


Fig. 132.1 WSN QoS architecture

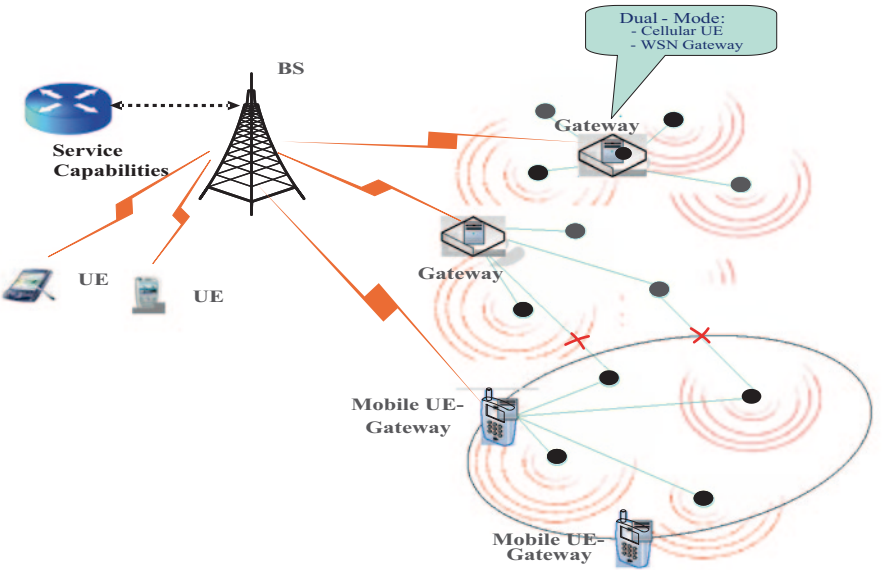


Fig. 132.2 System application scenario

Table 132.1 Applications and services requirements

Service applications	Requirements
e-Health care Monitoring vital signs Telemedicine points Remote diagnostics	Need to divide the transmission service types, emergency communication should have a high priority Guarantee different service having different delay and packet loss level Manage the data transmission takes place in applications Avoid access to a single channel resource
Metering: Industrial metering power Gas Water Heating	Monitored and controlled by a central server entity Have the ability to control metering device The network should be optimized to enable to transmit mass data simultaneously For the mobile UE can move, the sensor should be able to communicate with low mobility Be able to communicate with other connected object
Remote monitoring/control Environments surveillance Pollution and disaster surveillance	Fixed in remote areas for many years, low communication costs are extremely important Allowing detected data transmission during fixed time period Can dynamically adjust these time periods WSN can be transmitted in an on-demand mode Have the capability to change the working state of sensor nodes, e.g. activate or deactivate
Security Surveillance systems Theft/vandalism application	Make alarm if moving by unauthorized persons Security mechanisms (in order to avoid stealing or breaking) Monitoring the mobile information and alarm activate and deactivate

Table 132.1 continued

Service applications	Requirements
Tracking and tracing Traffic information Road tolling Road traffic optimisation	Sensor module be able to get its detailed location Sensor module should be able to communicate in a mobility mode Sensor module and its interfaces enable to survive and operate after an accident Have a high priority when an accident happens

1. Multiple service types: Different applications have different service requirements. Therefore, different types of sensor nodes raise challenges for QoS support. Generally, sensor nodes of monitoring security are more important than those of monitoring environment.
2. Packet criticality: The data packet of different monitor system reflects the criticality of the real application, which has different transmission priority. For instance, some important application types may require higher priority. QoS guarantee mechanisms can be used in the system transmission architecture.
3. Unbalanced load: In most applications of WSNs, load mainly focuses on the data traffic of sensor nodes transmitted to the sink nodes/gateway. Load balance scheme should be designed for unbalanced data traffic in different application.
4. Energy balance: In order to prolong the WSN lifetime, energy consumption should equably distribute among most sensor nodes, and so the energy of sensor nodes in certain area will not be drained out very soon.
5. Multiple gateways: There maybe exist multiple gateways in the convergent MCN and WSN to collect the detected data and transfer them to the backhaul networks. But, if the sensors and gateways are not uniformly distributed then it can cause few gateways to overload.

132.3 QoS Current Research for WSN

The existed research related to the QoS in WSNs mainly is divided into two categories: MAC algorithm improvement, routing algorithm improvement, and cross-layer solutions.

For WSNs, MAC plays an important role in the communication QoS, i.e. channel access, transmission delay and energy consumption. The existing MAC algorithms for WSNs mainly include S-MAC, T-MAC and B-MAC, which cannot provide QoS guarantee. An implicit prioritized access algorithm (I-EDF) is designed for WSNs, where cellular backbone network is adopted [18]. All nodes are synchronized and use the earliest deadline first (EDF) schedule algorithm for packet transmission to guarantee the QoS for different services. Watteyne propose a dual-mode MAC protocol, which solved the transmission delay problem for different applications [19]. As soon as collision occurs, protected mode is used to offer collision-free transmission. DMAC is an efficient algorithm to reduce the energy consumption

and transmission delay, which is designed for WSN data gathering [20]. Z-MAC is a hybrid MAC protocol, which can switch between CSMA and TDMA mode for WSN node transmission access [21].

Current research of routing algorithm improvement mainly includes the following algorithms, i.e. SAR (sequential assignment routing) [22], SPEED (A Stateless Protocol for Real-Time Communication) [23], et al. SAR is the first routing algorithm to guarantee the WSN transmission QoS. It considers the energy, service type and packet priority of WSN nodes. SPEED routing algorithm is proposed with providing end-to-end guarantee for real-time service [24]. Recently, another WSN QoS-aware algorithm is proposed, where real-time data traffic is mainly focused on the image or video sensors [25]. This proposed algorithm finds the least cost and energy efficient path to satisfy the service requirement. The QoS optimizing scheme is proposed by improving the sensor data transmission algorithm, where it uses the base station to broadcast the QoS information to each sensor node [26]. M. Perillo et al. provide QoS guarantee scheme by the hybrid optimization of scheduling and routing algorithms [27].

In the above algorithms, the sensor data is transmitted to the WSN sink unit by multi-hop mode. The end-to-end transmission delay may become large. Simultaneously, sensor nodes near to the sink unit can forward data more than the nodes at which far from the sink node and its energy maybe exhausted more quickly, which causes the disconnection for WSN. Besides, the data criticality should be considered and different priority levels should be assigned. And QoS concerns in WSNs should not be simply described as a single problem of transmission reliability. Other factors should be taken into account, such as latency, energy, and bandwidth.

132.4 Future QoS Research Points for WSNs

QoS control mechanisms should reduce the energy consumption, transmission delay and packet loss in data delivery. Further, the data transmission should have its QoS requirements according to its service types of application. The following are main research points of QoS support in the convergent WSN and MCN.

1. QoS-based data traffic model: It is very important to analyze how the QoS-constrained traffic while minimizing transmission delay and increasing the packets loss. Especially for the mass data in the UE-gateway, it needs resource allocation to guarantee the data transmission.
2. QoS-based services types: The type of detected data denotes the criticality of the application scenario. QoS scheme may be required to distinguish packet importance and establish a priority structure. Some applications may require a diverse mixture of sensors for monitoring. Thereby it needs to define different service type for sensor nodes, which makes WSN QoS requirements more challenging in the heterogeneous network environment.
3. QoS-supported adaptive transmission types: In the conventional method, the WSN data is transmitted to the sink node by multi-hop mode. In the convergent WSN and MCN, it is vital for how to choose the optimal UE-gateway for guar-

anteeing service data transmission of high priority and how to finish the sensor nodes data transmission after using the mobile UE-gateway.

4. QoS-supported adaptive load balance: When one gateway covers a large number of sensor nodes or high traffic and its neighbor UE-gateway has fewer nodes and lower traffic, this causes load unbalance and fast energy consumption of some sensor nodes. How to manage WSN load balance under the help of BS and UE-gateway has become research hot point.
5. QoS-supported adaptive UE-gateway access: After a UE-gateway enters the WSN coverage area, it may cause the gateway access (re)selection or even regrouping of the WSN nodes. How to make a balanced trade-off between the complexity, performance gain and energy consumption etc during the UE-gateway (re)selection and WSN regroup process is an important issue for further research.

132.5 Conclusion

MCN and WSN are transforming from heterogeneous to convergent to satisfy the increasing requirements of M2M communications. WSN QoS technical challenge still exists in M2M communications. In this paper, we analyze the system architecture for QoS guarantee scheme, where we list many applications requirements. Further, we analyze the transmission QoS challenges taken by the converged MCN and WSN. Finally, some research open issues are stated in order to stimulate more QoS research. We are convinced that the QoS in the MCN and WSN converged process should widely be researched in order to stimulate M2M application in the future, which can also promote the industry development of the Internet of things (IoT).

Acknowledgements This work was supported by the National Natural Science Foundation of China (No.61302113), Shanghai Science Foundation (No. 13ZR1440800) and Shanghai Rising-Star Program (14QB1404400), Shanghai Science and Technology Innovation Program (No.12511503300, 12DZ2250200, 14511101303).

References

1. Brazell J, Donoho L, Dexheimer J. M2M: the wireless revolution [M]. 1st ed. USA: TSTC Publishing; 2005. pp. 115–4.
2. Verdone R, Dardari D, Mazzini G, et al. Wireless sensor and actuator networks [M]. 2nd ed. London: Academic Press; 2010. pp. 86–95.
3. Puccinelli D, Haenggi M. Wireless sensor networks: applications and challenges of ubiquitous sensing [J]. IEEE Circuits Syst Mag. 2005;5(3):19–31.
4. Hu F, Kumar S. The integration of ad hoc sensor and cellular networks for multi-class data transmission [J]. Ad Hoc Netw. 2006;4(2):254–2.
5. Zheng J, Jamalipour A. Wireless sensor networks: a networking perspective [M]. 1st ed. New York: Wiley; 2009. pp. 125–8.

6. KAl-Omari SA, Sumari P. An overview of mobile ad hoc networks for the existing protocols and applications [J]. Int J Appl Gr Theory Wirel Ad Hoc Netw Sens Netw. 2010;2(1):87–110.
7. Zara H, Hussain FB. QoS in wireless multimedia sensor networks: a layered and cross-layered approach [J]. Wirel Pers Commun. 2014;75(1):729–57.
8. Crosby GV, Vafa F. Wireless sensor networks and LTE-A network convergence [C]. In: Proc. of the 38th Annual Conference on Local Computer Networks, LCN 2013, IEEE, Sydney, Australia, October; 2013 pp. 731–4.
9. Zhang J, Shan LH, Hu HL. Mobile cellular networks and wireless sensor net-works: toward convergence [J]. IEEE Commun Mag. 2012;50(3):164–9.
10. ETSI TS 102 689 V1.1.1. (2010–08) M2M service requirements.
11. ETSI TR 102 691 V1.1.1. (2010–05) Smart metering use cases.
12. Rec. ITU-T Y.2221. (2010–01) Requirements for support of ubiquitous sensor network (USN) applications and services in the NGN environment.
13. 3GPP TS 22.368 V10.2.0. (2010–09) Service requirements for machine type communications.
14. Bouckaert S, Poorter ED. Strategies and challenges for interconnecting wireless mesh and wireless sensor networks [J]. Wirel Pers Commun. 2010;53(3):443–3.
15. Chen DZ, Varshney PK. QoS support in wireless sensor networks: a survey [C]. In: Proc. of the International Conference on Wireless Networks (ICWN'04), CSREA Press, Las Vegas, USA, June; 2004, pp. 227–33.
16. Zhu J, Papavassiliou S, Yang J. Adaptive localized QoS-constrained data aggregation and processing in distributed sensor networks [J]. IEEE Trans Parallel Distrib Syst. 2006;17(9):923–3.
17. Alam MM, Mahbub M, Razzaque MA. Energy-aware QoS provisioning for wireless sensor networks: analysis and protocol [J]. J Commun Netw. 2009;11(4):390–5.
18. Caccamo M, Zhang LY, Sha L, et al. An implicit prioritized access protocol for wireless sensor networks [C]. In: Proc. 23rd IEEE Real-Time Systems Symposium, IEEE Austin, USA, December 3–5; 2002. pp. 39–48.
19. Watteyne T, Auge-Blum I, Ubeda S. Dual-mode real-time MAC protocol for wireless sensor networks: a validation/simulation approach [C]. In: Proc. International Conference on Integrated Internet Ad hoc and Sensor Networks (InterSense'06), IEEE, Nice, France, May 30–31; 2006. pp. 1–10.
20. Lu G, Krishnamachari B, Raghavendra CS. An adaptive energy efficient and low-latency MAC for data gathering in wireless sensor networks [C]. In: Proc. Int. Parallel Distrib. Process Symp., IEEE, Santa Fe, USA, Apr.; 2004. pp. 224–31.
21. Rhee I, Warrier A, Aia M, Min J. Z-MAC: a hybrid MAC for wireless sensor networks [C]. In: Proc. ACM Sensys., ACM, St Diego, USA, November 2005; 2005. pp. 90–101.
22. Sohrabi K, Gao J, Ailawadhi V, et al. Protocols for self-organization of a wireless sensor network [J]. IEEE Pers Commun. 2000;30(5):16–7.
23. Ye W, Heidemann J, Estrin D. Medium access control with coordinated adaptive sleeping for wireless sensor networks [J]. IEEE/ACM Trans Netw. 2004;12(3): 493–506.
24. He T, Stankovic JA, Lu CY. SPEED: a stateless protocol for real-time communication in sensor networks [C]. In: Proc. of the International conference on Distributed Computing Systems, IEEE, Providence, USA, May 2003; 2003. pp. 46–55.
25. Mahapatra A, Anand K, Agrawal DP. QoS and energy aware routing for real-time traffic in wireless sensor networks [J]. Comput Commun. 2008;29(4):437–5.
26. Singh J, Singh B, Chaudhary A. Ubiquity of mobile computing in wireless networks [J]. Int J Technol Eng Syst. 2009;1(1):1–4.
27. Perillo M, Heinzelman W. Providing application QoS through intelligent sensor management [C]. In: Proc. of the 1st IEEE International Workshop on Sensor Network Protocols and Applications (SNPA'03), IEEE, Anchorage, USA, May 2003; 2003. pp. 163–73.

Chapter 133
The Application and Experiment of High-Pressure Water Mist Fire Extinguishing System in the Tobacco Industry

Tinade Luo, Jing Qiao and Beihua Cong

Abstract The tobacco industry is an important industry in China's economic construction. Tobacco warehouses and archives, such as high- and low-voltage rooms, have an urgent need of secure, reliable, and energy-saving fire safety systems, in order to prevent all kinds of unexpected fires. This chapter analyzes the tobacco warehouses and archives, such as high- and low-voltage power distribution rooms with fire risk, which are able to illuminate the mechanism of water mist fire suppression systems and its features. We conducted a fire test in a real tobacco warehouse and checked the validity of the high-pressure water mist system to assure the safety of the tobacco warehouse.

Keywords Fire risk · Tobacco warehouse · File reference room · High-pressure water mist fire protection system

133.1 Introduction

During the "Eleventh Five-Year" the State Tobacco Monopoly Bureau conducted a large-scale integration of the national cigarette factories. Most of the old branches were closed while new factories under automated management were being constructed.

There are lots of regulations adapted for the production-storage mode of large foreign tobacco companies, combined with the automated production equipment and transforming the old tobacco factories into joint production factories. That is to combine the processing lines, expand the tobacco lines and cigarette envelopes, and elevate the library which include library materials and finished products' recipes.

T. Luo (✉) · J. Qiao · B. Cong
Tobacco Guangdong Industrial Co., Ltd., 510385 Guangzhou, China
e-mail: scltd@163.com

J. Qiao
e-mail: 22313335@qq.com

B. Cong
e-mail: congbeihua@tongtai-sh.com

W. Wang (ed.), *Proceedings of the Second International Conference on Mechatronics and Automatic Control,* Lecture Notes in Electrical Engineering 334,
DOI 10.1007/978-3-319-13707-0_133

In each building, facilities and equipment within the cigarette factory such as the boilers control room, compressed air machines, and air conditioners, even the smoke production equipment, are highly valued. For example, a desk M8 machine costs over 100 million Yuan.

Therefore, the production and storage areas of the cigarette factory are under risk to cause a big fire hazard [1]. In this chapter, the author gives a description in accordance with the participation in the transformation system operation and management of the cigarette factory. Meanwhile, the author gives an illustration of why the installation of high water mist fire extinguishing system is the first choice.

133.2 Water Mist System and Principle

In recent years, domestic and foreign institutions have invested a lot of manpower and resources to develop water mist fire suppression technology and products. The practice shows that the water mist fire extinguishing system, especially high-pressure water mist extinguishing system, expands the scope of the application of the traditional system with both the advantages of water spray and gas extinguishing [2].

133.2.1 Principles

It can also be defined by the National Fire Code of Dv0.99 that is less than 400 μm or 200 μm. Water mist fire extinguishing systems mainly rely on the endothermic vaporization cooling effect, which is to cut off oxygen by the extinguishing fires stifling effect. It is the basic mechanism of water mist fire suppression [3, 4, 5].

133.2.1.1 Vaporization Endothermic Cooling Effect

Since only one tenth the size of the water droplets of water spray, specific surface area expanded 10–100 times, its efficiency is 200–300 times [5]. Therefore, when the ambient temperature is rising, the water mist can quickly vaporize to absorb a lot of burning calories to reduce the fire temperature.

133.2.1.2 Isolated Choking Effect of Oxygen

With fine mist spraying into the fire, the water droplets absorb a lot of heat in the vaporization process, and the rapid expansion of volume can be enlarged 1700 times.

Table 133.1 Classification of water mist fire extinguishing system

System pipe network	Working pressure (MPa)	Type	System application
Single system Double fluid system	Low pressure(≤1.21) Pressure(1.21, 3.45), High pressure(≥3.45)	Open systems Close systems (Preaction, wets)	Overall situation Part systems

133.2.2 Classification

According to the classification, the system can be divided into three parts: high-pressure systems (>3.45 MPa), medium-pressure systems (1.21 MPa to ~3.45 MPa) and low-voltage systems (≤1.21 MPa). The system application can be divided into open and close systems (wet, dry, and preaction systems). Currently, the tobacco warehouses should use a precaution control system for safety; the wet control system is designed in this way (Table 133.1).

133.3 High-Pressure Water Mist and Water Spray Technology Are Relatively Fine Dry Powder Fire Extinguishing Systems

According to the statistics, the tobacco warehouse fire extinguishing system may be installed with high-pressure water mist system, water sprinkler system, and fine dry powder system. The performance of each of the three fire extinguishing systems is compared.

133.3.1 Fire effects

The water spray system has a water loss after a big fire. The super fine powder extinguishing system cannot uniformly cover the complete combustion surface, fire effects are inadequate [2], cannot sustain in fire. High-pressure water mist fire control systems are able to control the oxygen barrier stifling effect smoldering fires and can efficiently fight the fire without damaging the electrical equipment [4].

133.3.2 Environmental Impact

High-pressure water mist, which reduces the sedimentation through droplets, adsorbs the gas concentration of the fire area and reduces smoke toxicity.

133.3.3 Interface with Relevant Professional

The high-pressure water mist extinguishing system does not produce harmful gases after the fire, which also eliminates the need for separating smoke under fire mode, but needs to set up drainage measures in the pump equipment room.

133.3.4 Operations Management

The routine maintenance of the high-pressure water spray system is simple and costs little. And the pump just needs to be started once a month to check whether it works properly.

133.4 High-Pressure Water Mist Extinguishing System Experiments

133.4.1 Simulation High Rack Warehouse

The size of the simulation overhead warehouse is 12.0 m (L) × 12.0 (W) × 8.5m (H). At the top of the warehouse, the closed nozzle set $K = 2.04$, in each of the two shelves, the closed nozzle set $K = 1.25$. The closed nozzle response temperature is 57 ℃, working pressure is 10 MPa, response time index (RTI) is <20, sprinkler heads spaced no more than 3.0 m, as shown in Fig. 133.1.

The high-pressure water mist system to extinguish fires in the process elevated the warehouse shown in Fig. 133.2. From ignition to complete extinguishment, it took a total of 1 min 32 s, its closed nozzle response time was only 19 s. After the water mist system startup, the fire temperaturelowered quickly and effectively inhibited the fire achieving good results. After the success of the water mist fire suppression, the fire area remains small, the water loss and the impact on tobacco are small as well.

133.4.1.1 Analog Single Warehouse

The simulated single warehouse size 12.0 m (L) × 12.0 (W) × 5.6m (H). At the top of the warehouse, the closed nozzle setting $K = 2.04$. Closed nozzle response temperature is 57 ℃, working pressure is 10 MPa, and RTI <20. Maximum nozzle spacing is 3.5 m, as shown in Fig. 133.3.

Figure 133.4 shows a set of test images. Two minutes after ignition with diesel, the glass bulb of high-pressure water mist closed nozzle ruptures and release of water mist begins. At this moment, we should cover the oil basin fire with a cover plate. After 50-second mist spray, the flame on the smoke box has been restrained,

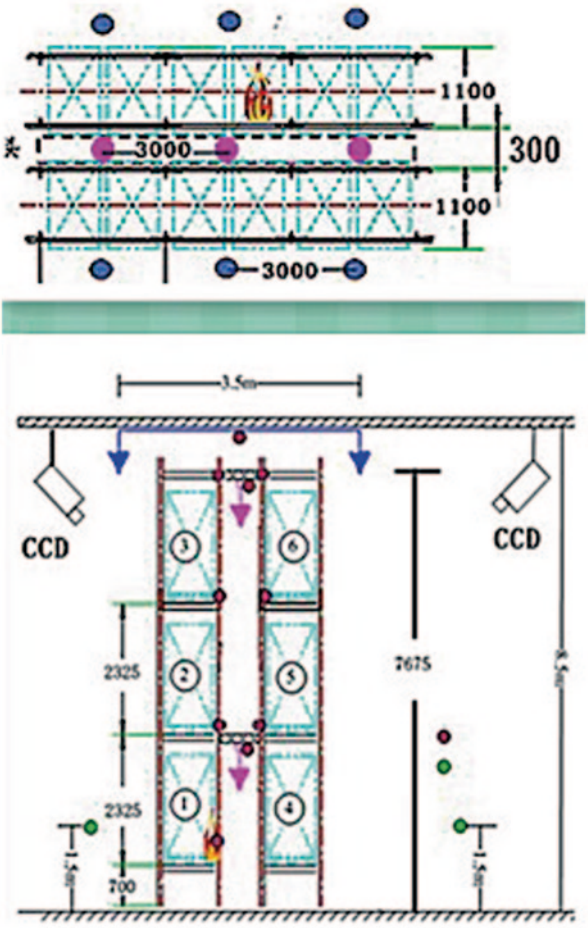


Fig. 133.1 Simulated test overhead warehouse layout (mm)

with only a few sporadic small flames left. 2 minutes and 33 seconds' spray later, the small flames haven't been completely extinguished, with smoldering fires in the smoke box. Then the control valve at the top of the water mist nozzle is to be closed, and the water mist gun begins to work. 1'29" later, extinguish small fires completely and continue to spray at regular intervals for 1 min, then all the smoldering fires will be put out.

133.4.1.2 Shi Ma Warehouse Fire Test

The size of the test storage is 120.0 m (L) × 60.0 (W) × 4.0 m (H), provided at the top of the warehouse, library closed nozzle $K = 1.25$. Closed nozzle response temperature is 57℃, working pressure is 10 MPa, and RTI < 20. Maximum nozzle spacing is 3.0 m. The test arrangement is shown in Fig. 133.5. The packing-case is ignited by

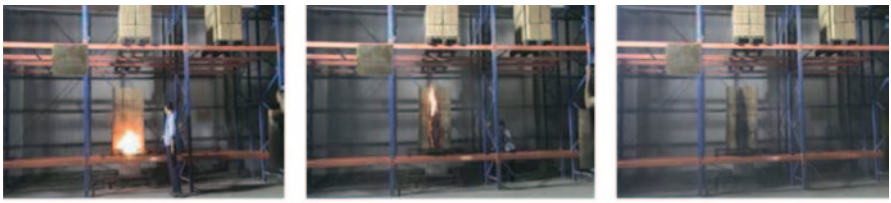


Fig. 133.2 Analog overhead warehouse fire test procedure. **a** Ignition. **b** Extinguishing process. **c** Extinguishing completed

Fig. 133.3 Single warehouse layout test

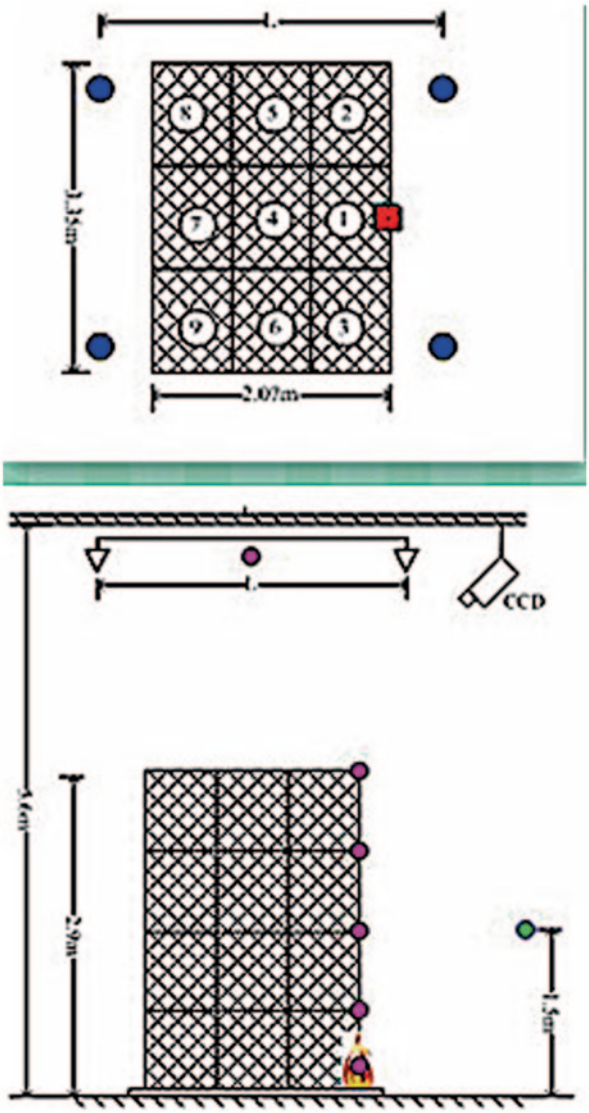


Fig. 133.4 Figure simulation single warehouse fire. **a** Ignition. **b** Extinguishing process. **c** Extinguishing completed

petrol in the test. The size of the ignited oil disc is 0.5 m (L) × 0.5 m (W) × 0.2 m (H), and the amount of ignited gasoline is 500 ml. Two ignition points are set, which are respectively located below and in the middle of the nozzle. In addition, a high-pressure water mist spray gun is set on the spot, the working pressure of which is 10 MPa.

Figure 133.6 shows the spray water mist fire extinguishing process. After ignition, after 24 s the closed glass bulb sprinkler bursts and sprays the fire, the flames are extinguished after 1 min and 51 s, and the discharge of water mist spray is stopped. Three minutes and 34 s after the flame resurgence, the water mist fire extinguishing system reboots and uses the gun to assist the fire, and after 28 s, the flame is completely extinguished.

133.4.2 Typical Examples of Introduction

133.4.2.1 Protected Area

This project uses the high-pressure water mist fire suppression system to protect a high-bay warehouse overhead warehouse. Protection of an area of 2448 m^2, the number of sprinklers is 576, using DN32 mm preaction valve box system, keeping the object of the warehouse shelves and the top layer.

133.4.2.2 System Design Type

The high-pressure water mist system by system action is divided into open and closed systems, including wet systems and preaction systems. The project is to improve the reliability of the system to prevent accidental spraying, using the high-pressure water mist preaction system.

133.4.2.3 Design Parameters

Roof closed nozzle is using $K = 2.04$, the inner shelf closed nozzle is using $K = 1.25$. Shelves within the nozzle which is layered set above the nozzle increases the collector hood. The nozzle design flow equation is:

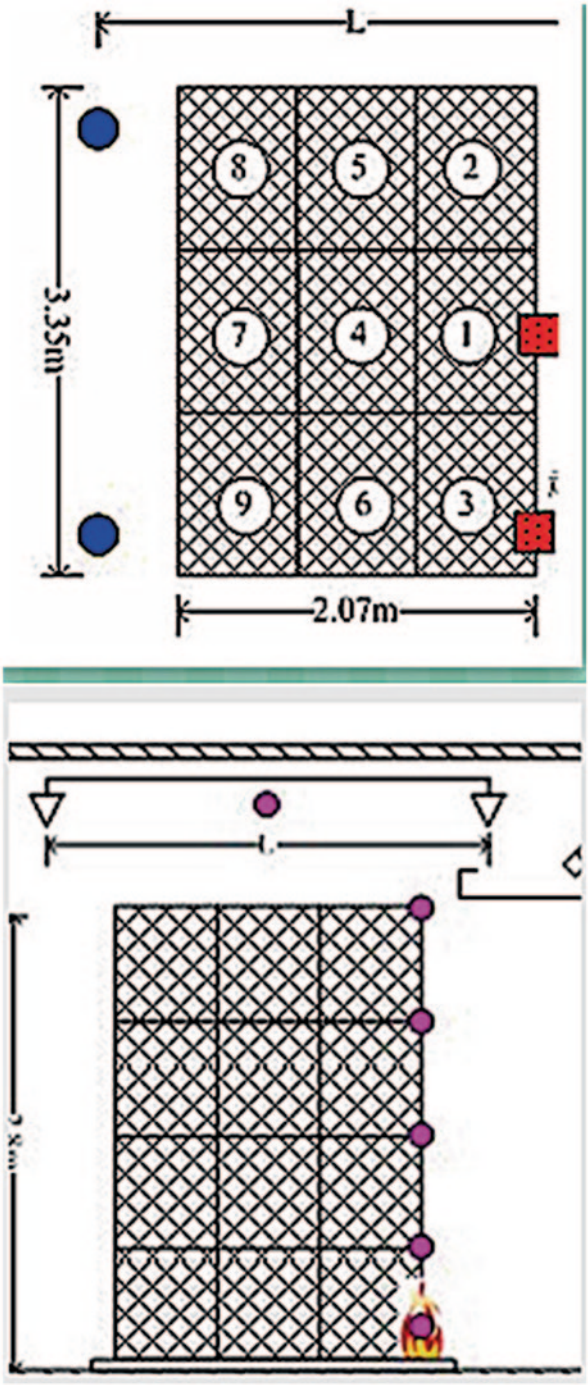


Fig. 133.5 Shi ma fighting tobacco warehouse layout entities

Fig. 133.6 Shi ma tobacco warehouse fire process diagram. **a** Ignition. **b** Fire started. **c** Fire inhibited. **d** Flame goes. **e** Recrudescence. **f** Fire started. **g** Gun fire. **h** Flame goes

$$q = K\sqrt{10P}. \tag{133.1}$$

Wherein the nozzle flow parameters K and P represent the pressure in MPa, the start nozzle temperature is 57℃, continuous spray time is no less than 30 min, and the minimum working pressure nozzle $P = 10$ MPa.

The installed closed nozzle spacing is not greater than 3 m, not less than 2 m, and the wall is not greater than the distance of 1.5 m. Therefore, the sprinkler heads protected areas are determined, which is shown in Fig. 133.7–133.8.

Calculated flow system:

$$Q = 1.1{*}q{*}N. \tag{133.2}$$

Wherein q is the flow rate nozzles, N is the number of nozzles and Q is the units of L/min. The system flow is simultaneously opening 14 nozzle flows, calculation is based on all the shelves sprinkler nozzle flows and the role of the open area of 140 m^2 which leads to the system flow of 413 L/min. System design working pressure is according to the most negative point of minimum working pressure of 10 MPa to calculate the nozzle. The pipe network hydraulic calculation uses the Darcy–Weisbach equation, which derived $H = 13$ MPa.

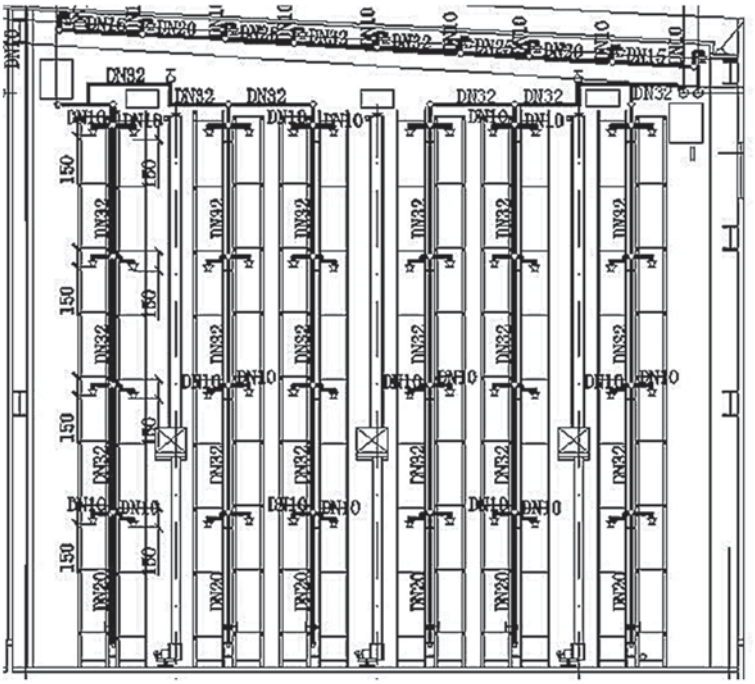


Fig. 133.7 The head portion disposed within the plane of the shelf (mm)

133.5 Conclusion

The high-pressure water mist fire extinguishing system is a new and efficient system, a saving and investing and environmentally friendly fire extinguishing system. The fire test proves the good effect of the high-pressure water mist. Studies show that the high-pressure water mist has good smoke and dust effects and improves visibility, which will help firefighters to enter the reservoir disposal. It can also reduce the temperature of the treasury to reduce the indirect loss caused by fire.

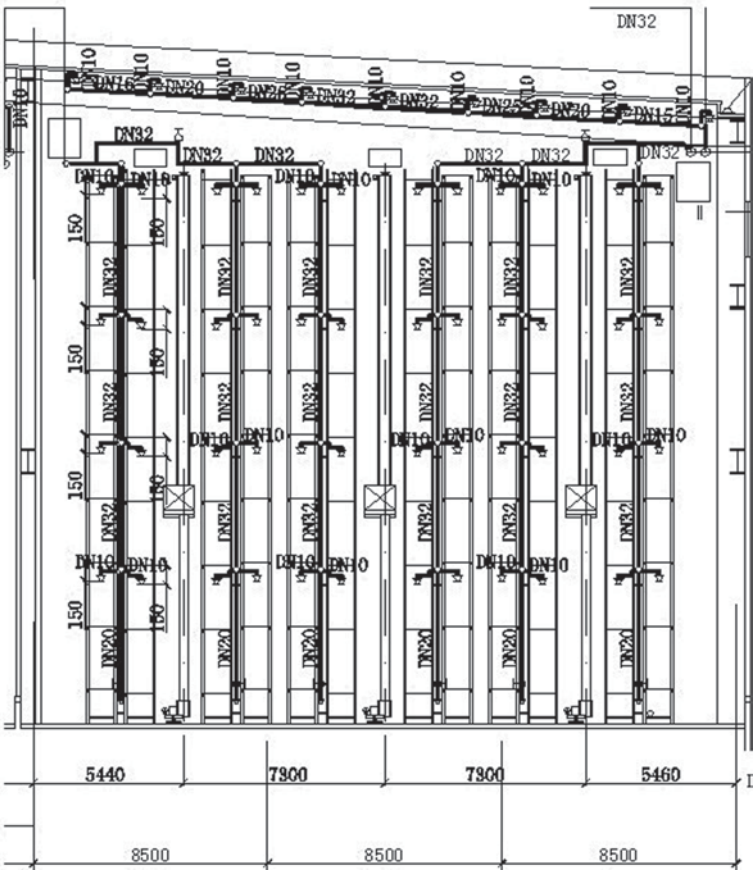


Fig. 133.8 Shelves and roof cross-section of the nozzle arrangement (mm)

References

1. Zhang H. Danger and fire countermeasures tobacco production, storage fire. J Chinese People's Armed Police Force Acad. 2008;6(2):47–9. (in Chinese).
2. Jing J. Applicability analysis device for use in high-bay warehouse cigarette hanging gas supersonic fixed automatic fire extinguishing powder. Chinese Tob. 2009;6(4):8–11. (in Chinese).
3. Guangxuan L, Xin H, Cong B, Jun Q, Liu J, Wang X. Progress in water mist fire suppression technology. J University Sci Technol China. 2006;36(1):9–19.
4. Aetna Engineering, Tongji prevention research center. Test report high pressure water mist in the tobacco warehouse of.3(2); 2009. (in Chinese).
5. Liu C. High-pressure water mist fire extinguishing system in Guangzhou Zhujiang New City automated passenger transport system application. Anhui Build. 2012;6(2):173–174. (in Chinese).